Alternative Information Signs: Evaluation of Driver Comprehension and Visual Attention

Justin M. Neill[1]; David S. Hurwitz, A.M.ASCE[2]; and Michael J. Olsen, A.M.ASCE[3]

Abstract: The effectiveness of a traffic sign is collectively influenced by the sign's understandability, legibility distance, glance legibility, and learnability; however, understandability has been repeatedly identified as one of the most important effectiveness measures. This study contributes to best practices for evaluating traffic sign understandability by demonstrating and comparing a variety of online survey questions and immersive driving simulation tasks. These techniques were applied to assess the understandability of five alternative tourist information (TI) signs in Oregon. Several TI sign alternatives were first tested in an online survey with 142 participants, followed by more authentic testing of 42 participants in the Oregon State University Driving Simulator. Sign INFO was correctly understood by 95.7% of the driving simulator subjects. Similar results were obtained for the other testing modalities. Notably, the understandability of TI sign alternatives, including versions with (75.4%) and without (72.8%) a circular border, did not score as high despite their prevalent usage in other contexts. However, the Sign i comprehension rates increased dramatically over those from an earlier study, indicating that periodic review of sign comprehension rates may be needed to reflect changes in understandability with time. **DOI: 10.1061/(ASCE)TE.1943-5436.0000807.**

Author keywords: Traffic signs; Tourist information signs; Driver understandability; Driver comprehension; Driving simulator.

Introduction

The gross domestic product of the travel industry in the state of Oregon was $3.5 billion in 2013, placing it among the three largest export-oriented industries in the state (Dean Runyan Associates 2014). The travel industry also has a significant secondary effect on employment in Oregon. In 2013, the respending of travel-related revenues by businesses supported 42,300 additional jobs outside of the travel industry (Dean Runyan Associates 2014). Intuitive access to visitor information centers is a vital contributing factor to the potential economic impacts of tourism in Oregon and elsewhere. Tourist information signs are intended to direct roadway users to nearby tourist information centers (FHWA 2009). However, for these signs to work effectively, they need to be easily interpreted and understood by a wide variety of visitors.

Because of the economic importance of domestic and international tourism and the wide variability in sign display, there is a significant interest in the comprehension rate of tourist information signs. It has been widely proven in the literature that symbolic messages elicit a higher comprehension rate than a text-based message because of its language-independent nature and the minimal space needed to deliver the intended message (Walker et al. 1965) as well as greater legibility distances (Babbitt Kline et al. 1990). Paniati (1988) and Katz et al. (2012), among numerous others, demonstrated that high rates of driver comprehension could be achieved by replacing test-based sign messages with symbols. Earlier research was conducted on information sign comprehension using surveys and on sign legibility using a sign simulator that displays a single image at a time on a two-dimensional (2D) projector (Katz et al. 2008).

In contrast, the study described in this paper is distinct from the work of Katz et al. (2008) because driver comprehension was tested with open-ended survey questions, rating survey questions, and open-ended interview questions and an evaluation of visual attention while subjects were engaged in a simulated driving task facilitated by a high-fidelity full-scale driving simulator. Further, this experimental design not only provided information about the comprehension of the information sign but also informs best practices for determining sign comprehension. Specifically, this study represents a unique contribution because: (1) four different modalities were evaluated to determine whether comprehension results would vary based on the selected modality—often comprehension tests include only one or two modalities; (2) the 6-year offset between the Katz et al. (2008) study and the results of this study provide a comparison of comprehension rates over time for signs; and (3) this study examines several additional information sign alternatives not considered in the prior study.

Background

To strengthen the argument for the aforementioned research focus and to set the stage for the experimental design, a brief review of the literature relevant to tourist information signs and traffic sign comprehension is provided. The findings are presented in four focus areas: tourist information sign standards and contexts, relevant research on traffic signs in general and Sign i in particular, and sign comprehension testing methods.

Tourist Information Sign Standards

The information message (which is intended to inform visitors that information is available to them) is used in a wide variety of

[1]Graduate Research Assistant, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331. E-mail: neillju@onid.oregonstate.edu

[2]Associate Professor, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331 (corresponding author). E-mail: david.hurwitz@oregonstate.edu

[3]Associate Professor, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331. E-mail: michael.olsen@oregonstate.edu

Note. This manuscript was submitted on July 18, 2014; approved on July 16, 2015; published online on October 5, 2015. Discussion period open until March 5, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, © ASCE, ISSN 0733-947X.

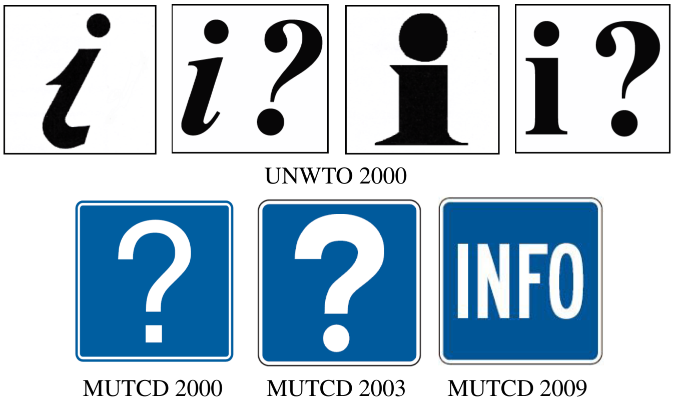


Fig. 1. UNWTO recommended and MUTCD required tourist information signs (FHWA 2000, 2003, 2009)

contexts, including tourism, shopping, software help, and general information. The tourist information sign [D9-10] is defined as a general service guide sign by the manual on uniform traffic control devices (MUTCD), which sets standards for traffic control devices, including signage, in the United States (FHWA 2009). Internationally, alternative tourist information signs have been recommended and adopted. For example, on November 30, 2000, the United Nations World Tourist Organization (UNWTO) executive council adopted four possible sign symbols (Fig. 1) to indicate the location of an information center (UNWTO 2000). Despite their differences, these symbols have one commonality: the use of a lowercase letter "i."

While the MUTCD and the UNWTO have both proposed information sign standards, other symbols have been used in a variety of transportation- and non-transportation-related contexts. Hence, it is important to observe other symbols that have been used to indicate locations where information is available and to observe what messages the symbols in Fig. 1 have been used to represent.

Information Signs and Symbols in Different Contexts

The numerous contexts of information signs and symbols necessitate additional investigation to inform transportation- and travel-related applications. The contexts range from smart phones to Internet-enabled mapping sites to airports. The most commonly recommended approach, based on a review of the literature, is the use of a lowercase letter i or ? as the information symbol. While a variety of the Sign i are used in several different contexts, it should be noted that many of these presentations include a circular background surrounding the i. Another common approach is the use of a Sign ? to indicate the availability of information with and without the inclusion of a circle as an element of the symbol. The i and ? symbols have been used in a variety of contexts and typically have been used to indicate that information is available (Fig. 2). Many of these symbols have also been used internationally to communicate meanings.

Relevant Traffic Sign Research

It has been postulated that certain inherent characteristics are required for traffic signs to perform effectively. Numerous studies have been conducted to describe these characteristics, including three seminal studies by Dewar (1988), Smiley et al. (1998), and Castro et al. (2004) (Fig. 3). Fig. 3 organizes critical sign characteristics (terms in bold) and their definitions (phrases not bolded) as determined in each of the three seminal studies. The arrows connect the critical sign characteristics with similar definitions or different terms between the different studies. Some characteristics, such as understandability, appear in all three studies, but under different terms (understandability, comprehension, and understand) with slightly different definitions, while other sign characteristics, such as learnability, appear in only one study.

By interviewing participants with expertise in traffic control devices and traffic engineering, Dewar rank-ordered six traffic sign design criteria. Participants determined that conspicuity followed by understandability were the most important criteria for information signs (1988).

Fig. 2. Example information signs and symbols from alternative contexts (images by Michael J. Olsen; image courtesy of Pixabay)

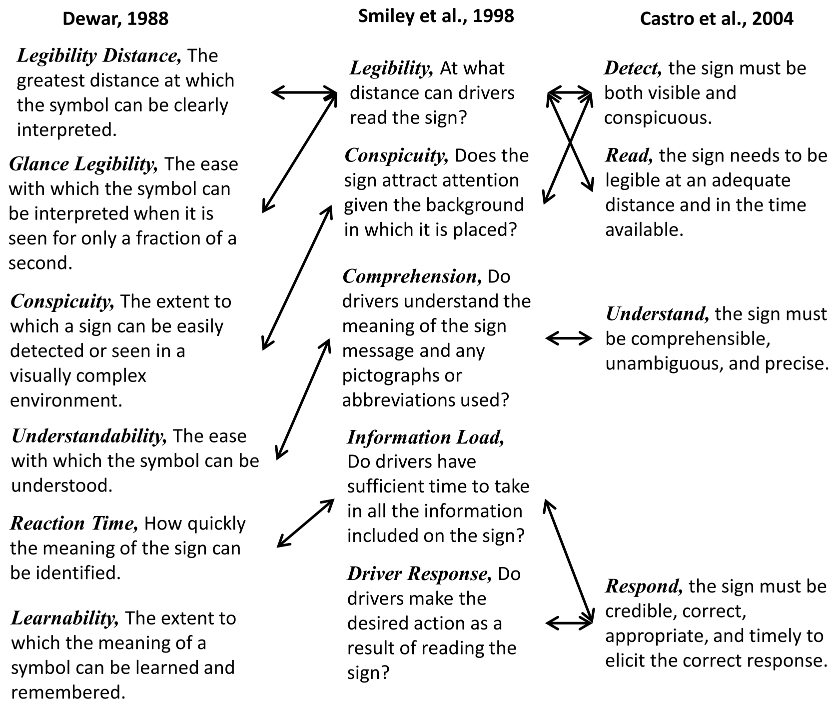


Fig. 3. Critical sign characteristics and relations over time

In 1998, while designing a new tourist signing system for the Ministry of Transportation in Ontario, Canada, Smiley et al. arrived at the same top five ranked criteria determined by Dewar in 1988.

Castro et al. (2004) proposed that four sequential stages exist when a road user interacts with a traffic sign, and each stage has a key consideration. In each of the four stages—detect, read, understand, and respond, as defined in Fig. 3—the driver uses some aspect of the sign to accomplish the necessary interaction (Castro et al. 2004).

Sign i Research

The i symbol was tested for comprehension and legibility by Katz et al. (2008), who compared it with the INFO word message and the ? symbol. The comprehension research was conducted using open-ended questions followed by multiple-choice questions in a survey. Katz et al. found that 56% of the subjects understood the correct meaning of the i symbol compared to 68% with the ? symbol and 96% with the INFO message when presented with the open-ended test. The multiple-choice questions included four alternatives: use caution, wireless Internet availability, medical assistance, and traveler information. The multiple-choice questions resulted in 76, 92, and 95% correct answers for the i, ?, and INFO messages, respectively. The most significant risk for the transferability of multiple-choice traffic sign comprehension surveys is the quality and plausibility of distractor questions (Wolff and Wogalter 1998). As such, the observation that several of the distractors were not selected for some of the sign alternatives may bring into question their plausibility. Additionally, Katz et al. found that, as tested, Signs i and ? had statistically greater legibility distances than Sign INFO.

Methodology

The methods employed included an online survey to determine general public understanding and preference for information signs and a human-factors assessment of actual responses to signs in a driving simulator. Each of these methods had two distinct tests, resulting in a total of four methods tested. The online survey produced data from both open-ended and rating comprehension questions. Data from both parts of the online survey were analyzed across subject demographics. The driving simulator data provided measurements of visual attention and accuracy of verbal responses.

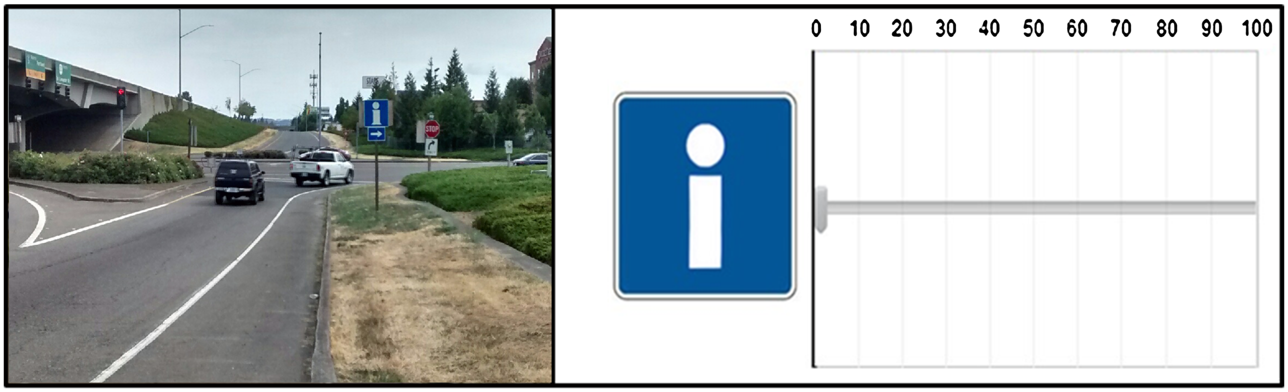
Research Hypotheses

The primary objective of this research is to determine which tourist information sign has the highest level of understandability by evaluating the comprehension and glance patterns between sign alternatives and test methods. Specifically, the four null hypotheses examined were as follows:

1. There is no difference in driver comprehension between each sign alternative;
2. There is no difference in a driver's glance patterns or fixation points between each sign alternative;
3. There is no difference in a driver's glance patterns or fixation points between correct, partially correct, and incorrect responses; and
4. There is no difference in driver comprehension between each sign alternative in the online survey and in the driving simulator.

Online Survey

The online survey consisted of demographic questions, open-ended sign comprehension questions [Fig. 4(a)], and rating questions [Fig. 4(b)]. Fig. 4(a) is a digital photo of an information sign location in Oregon with a digitally imposed image of Sign i. Immediately below this image in the survey, participants were directed to explain in their own words the meaning of the particular information sign alternative presented. Open-ended questions were used because they reflect the recommended practice of the American National Standards Institute (ANSI) Z535.3 (ANSI 2011); in



(a) (b)

Fig. 4. Example images of (a) comprehension and (b) rating (image by authors) questions from online survey for Sign I

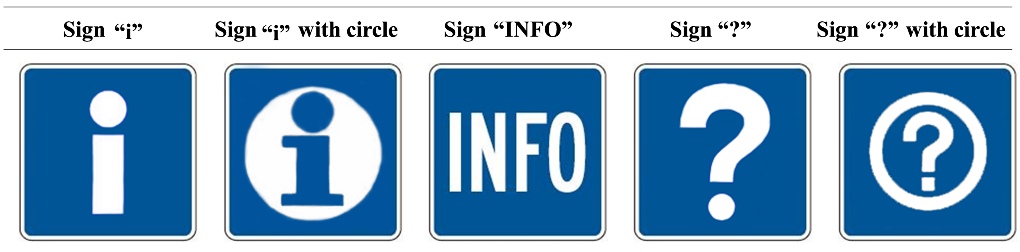


Fig. 5. Alternative signs tested (FHWA 2009)

addition, such questions make it possible to avoid the impact of poorly selected distractors (Wolff et al. 1998).The rating questions were included owing to the findings of Zwaga (1989), which showed that subject estimates of population comprehension can reliably be used as an early indicator for the usefulness of a symbol.

Subjects who were younger than 18 years or older than 75 years old or who had not been a licensed driver for more than 1 year were excluded from the study.

For the open-ended comprehension questions, subjects were presented with five alternative information signs (Fig. 5) on one of two different authentic Oregon roadway backgrounds to provide an authentic context because the presentation of signs in a realistic context was shown to greatly facilitate comprehension (Wolff et al. 1998). The order of sign alternatives and the backgrounds presented to each subject were randomized.

As seen in Fig. 5, three basic sign types were selected for the survey based on those found in a literature review: Signs i, ?, and INFO. A slight variation was included for both Signs i and ?. In addition to the symbol, a circle around the symbol was also considered. The inclusion of no more than one variation of an individual symbol (i.e., Sign i and Sign i with a circle) for each subject is consistent with the ANSI Z535.3 recommendation.

Prior to beginning the rating questions [Fig. 4(b)], subjects were presented with the following description of tourist information centers, provided by Travel Oregon: "Tourist Information Centers provide brochures, directions, and information about the surrounding area. This information includes local and regional activities and tourist attractions, as well as information about local restaurants and lodging." Subjects were then asked to "Select the percentage of the population you think will understand the following signs to represent a Tourist Information Center." As with the comprehension questions, all five of the sign alternatives (Table 1) were presented in a random order to each subject.

Driving Simulator Study

During the driving simulator experiment, the subjects' comprehension of alternative tourist information signs was assessed while they were engaged in a simulated driving task. The same signs were tested in the driving simulator experiment as those that were tested in the online survey, except for Sign ?. Sign ? was removed because it consistently generated the lowest comprehension rates. In a brief follow-up, subjects were given an online survey to rate the four signs they encountered during the driving simulator experiment, according to the percentage of the drivers in the United States that would correctly understand the sign.

Driving Simulator

The Oregon State University (OSU) driving simulator is a high-fidelity simulator, consisting of a full 2009 Ford Fusion cab mounted on top of a pitch motion system [Figs. 6(a) and 5(b)].

Table 1. Open-Ended Test t-Test p-Values

Sign alternative	Sign i	Sign i with circle	Sign INFO	Sign ?	Sign ? with circle
Sign i	1.000	—	—	—	—
Sign i with circle	**0.070**	1.000	—	—	—
Sign INFO	**<0.001**	**<0.001**	1.000	—	—
Sign ?	0.684	**0.029**	**<0.001**	1.000	—
Sign ? with circle	0.365	0.362	**<0.001**	0.238	1.000

Note: Bolded values are statistically significant ($p < 0.05$).

J. Transp. Eng., 2016, 142(1): 04015036

(a) (b) (c)

Fig. 6. Views from (a) outside OSU driving simulator; (b) inside OSU driving simulator; (c) subject wearing eye-tracking device (image by authors)

The pitch motion system accurately models acceleration and braking. Three projectors produce a 180° front view, and a fourth projector displays a rear image for the driver's center mirror. The two side mirrors have liquid crystal displays. The vehicle cab instruments are fully functional and include a steering control loading system to accurately represent steering torques based on vehicle speed and steering angle. The simulator software can record performance measures such as speed, position, braking, and acceleration at a sampling rate of 60 Hz.

As can be seen in Figs. 6(a and b), the driving simulator provides an immersive, built environment and an authentic driving task that allows individual variables to be examined in isolation while controlling for confounding factors. The human-factors assessment was performed in the driving simulator with an Applied Science Laboratories (ASL) (Bedford, Massachusetts) mobile eye tracking system [Fig. 6(c)] and think-aloud interviews.

Eye Tracking

Eye-tracking data were collected using the Mobile Eye-XG platform from ASL [Fig. 5(c)]. The advanced Mobile Eye-XG allows the subject to have unconstrained eye movement and unconstrained head movement, generating a sampling rate of 30 Hz and with an accuracy of 0.5–1.0°. The subject's gaze is calculated based on the correlation between the subject's pupil position and the reflection of three infrared lights on the eyeball. Eye movement consists of fixations and saccades. Fixations occur when subjects focus on a point in their visual field for a short period of time, and saccades occur when the eye moves from one point to another. The Mobile Eye-XG system records a fixation when a subject's eyes have paused in a certain position for more than 100 ms, a commonly accepted value for fixations in the context of driving (Konstantopoulous et al. 2010; Marnell et al. 2013; Hurwitz et al. 2014). Saccades are not recorded directly but calculated based on the dwell time between fixations. However, in this paper, driver saccades were not analyzed.

Scenario Layout

The scenarios presented in the driving simulator were modeled after realistic presentations of tourist information signs in Oregon. The subjects were exposed to tourist information signs in two contexts: first, on a freeway exit, and second, at an intersection of local roads (Fig. 7). The four sign alternatives that yielded the best results in the online survey were selected as the signs tested in the driving simulator.

The route taken by the subjects included traveling northbound along a freeway, departing the freeway by an exit ramp, and then turning right twice along local roads. Before the test, each subject was instructed to take the first exit and then make a right turn onto the local road, followed by a right at the final intersection. Each subject drove through the environment a total of four times. During the first two drives, the subject was shown each of the four signs in one of the two sign positions shown in Fig. 7. During the second two drives, the subject drove through the environment again with

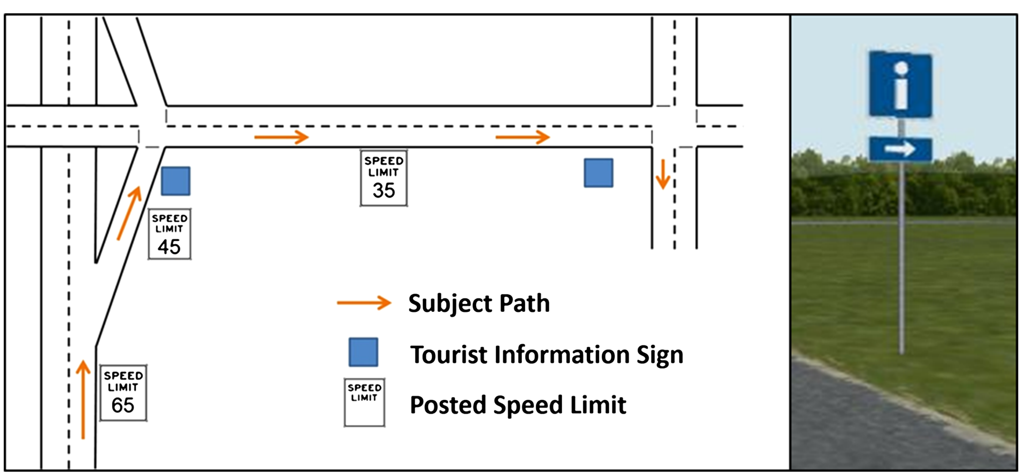


Fig. 7. Driving simulator track (not to scale) and example Sign i

the signs displayed in a different order. Throughout each drive, data were collected on the subject's lane position and speed. On the last two drives, as the subject approached each sign, they were asked to describe the meaning of the sign while they continued to drive through the environment.

A variety of equipment was used to record and track their responses. Verbal responses were recorded on a Zoom H2n Handy Recorder. The subjects wore the ASL Mobile Eye XG equipment through each of the four drives to record the visual attention of the subjects. Specifically, the fixations of the subject were measured to find the total number of dwells and the average dwell duration on each alternative sign. A single dwell was calculated as the sum of multiple, uninterrupted fixations on a single area. Fixations were calculated using ASL Mobile Eye postprocessing software.

Results and Analysis

This section describes the qualitative and quantitative data collected from the online survey and driving simulator experiment, the data reduction procedure, and the statistical methods used to analyze the data.

Online Survey

A total of 142 (68 male and 74 female) subjects responded to the online survey. The sample size complies with the ANSI Z535.3 recommendation for testing a minimum of 50 subjects for sign comprehension. The subjects' ages ranged from 19 to 73 years (average age of 34.3 years). The highest education level attained by subjects ranged from a high school diploma (6 subjects, or 4.2%) to a Ph.D. (3 subjects, or 2.1%), with the most common education level being a 4-year degree (56 subjects, or 39.4%). Kilometers driven in the previous year ranged from 0 to 8,047 km (5,000 mi) (29 subjects, or 20.4%) to more than 32,187 km (20,000 mi) (10 subjects, or 7.04%), with the most common amount of driving being between 8,047 and 16,093 km (5,000 and 10,000 mi) (33.8%). When compared to Oregon Department of Motor Vehicles records, the subject demographics are representative of the population of Oregon drivers.

Open-Ended Survey Questions

The 142 responses to the open-ended questions were classified independently by 5 researchers as 1 if correct, 0.5 if partially correct, and 0 if incorrect. The resulting data constituted discrete panel data. The use of multiple reviewers, all familiar with the symbol messages, is supported by previously conducted research (Wolff and Wogalter 1998). Open-ended responses were defined as correct if the subject demonstrated an understanding that the sign indicated that an information center was available nearby that could provide local information related to tourist activities. If a subject only demonstrated a partial understanding, the response was defined as partially correct. Examples of partially correct responses to the open-ended comprehension questions included "office of tourism and travel," "a location where you could only get route information," and "ask your questions here" without specifying the type of question being asked. For the question mark signs, partially correct answers included responses where the subject understood that the sign was directing them to a location where they could ask questions but were not clear with respect to the context of the questions. If a subject did not demonstrate an understanding, the response was defined as incorrect. To insure proper interrater reliability, any individual item that was not consistently scored by all five researchers was flagged. Those items were reexamined and discussed by the researchers until a consensus was reached.

In addition to the comprehension score, critical confusions were identified when the comprehension of the sign was dramatically different from the intended comprehension (ANSI Z535.3). Multiple comparisons were made with the results, including differences between gender, age, highest level of education completed, the number of miles driven in the previous year, the frequency of recreation or pleasure travel, the order that the signs were displayed, and whether the symbolic signs contained circular borders. The generated p-values were adjusted for the multiple comparisons through the Benjamini and Yekutieli (2001) adjustment.

Multiple recurring, wrong answers were given by the respondents. The most common incorrect interpretations were that the "i" symbols indicated pedestrians and that the blue background indicated a hospital sign. Misinterpretations of the signs indicating that a hospital was nearby were marked as critical confusions because of the potential to misguide someone in an emergency.

An ANCOVA test was used to test for differences in the means when considering the factors collected, followed by t-tests if a significant difference was found (Ramsey and Schafer 2013).

Models were created to determine which factors were important in driver comprehension. First, a full model, Eq. (1), was created by including all factors as additive variables:

$$\begin{aligned}\text{Mean Comprehension Score} &= \beta_0 + \beta_1 \times \text{Sign Type} + \beta_2 \times \text{Age} + \beta_3 \times \text{Education} \\ &\quad + \beta_4 \times \text{Sign Order} + \beta_5 \times \text{Years Licensed} \\ &\quad + \beta_6 \times \text{Recreation Travel Frequency} \\ &\quad + \beta_7 \times \text{Miles Driven in Previous Year} + \beta_8 \times \text{Language}\,e \end{aligned} \tag{1}$$

A reduced model was then found by comparing the full model with reduced versions until only significant variables remained, Eq. (2):

$$\text{Mean Comprehension Score} = \beta_0 + \beta_1 \times \text{Sign Type} + \beta_2 \times \text{Miles Driven in Previous Year} \tag{2}$$

The reduced model that emerged showed a significant impact of the sign type and of the miles driven in the previous year ($p < 0.001$ and $p = 0.010$, respectively). Then, t-tests were performed on the sign alternatives to determine which signs differed from the others; the results of these tests are shown in Table 1.

The statistical results presented in Table 1 show that Sign INFOwas comprehended better than all other sign alternatives. The Sign i with a circle alternative also outperformed Sign ? alternative.

Out of the eight driver demographics considered, only the miles that the subject drove in the previous year emerged as significant. The average comprehension score and standard deviation for each grouping of the kilometers driven in the previous year was 0.76 (0.42) for 0–8,047 km (0–5,000 mi), 0.85 (0.36) for 8,047–16,093 km (5,000–10,000 mi), 0.86 (0.35) for 16,093–24,140 km (10,000–15,000 mi), 0.88 (0.33) for 24,140–32,187 km (15,000–20,000 mi), and 0.72 (0.44) for more than 32,187 km (20,000 mi). Generally, the comprehension rate rises as the number of kilometers driven in the previous year rises, with the exception of the group who drove more than 32,187 km (20,000 mi) in the previous year.

Rating Task Survey Questions

Panel data with a continuous dependent variable were generated from the rating task when subjects were asked to rate each sign

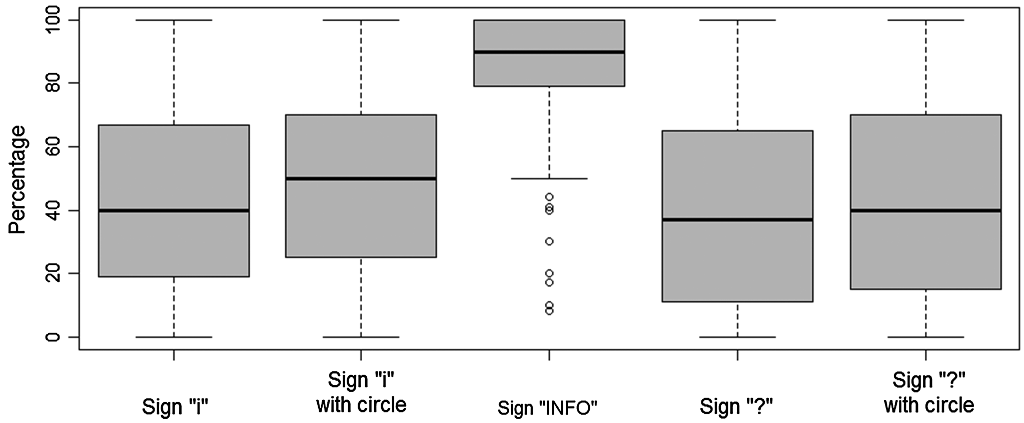


Fig. 8. Sign comprehension scores from rating task

with the percentage of the United States population that would correctly understand each of five information sign alternatives. Ten outliers, which are defined as data points outside 1.5 times the interquartile range above the upper quartile and below the lower quartile, were found and removed before the analysis was conducted. Box and whisker plots were created for each of the signs in the rating task (Fig. 8) and illustrate the comparison between each sign alternative. The Sign INFO alternative was consistently rated best compared to the other tested alternatives.

Both random- and fixed-effects models were considered to fit the online survey rating task panel data. A two-way model was chosen to account for the bias that may have occurred owing to subjects' making multiple observations. The Hausman test was conducted on the additive model, and it was found that the random-effects model fit the data better ($p > 0.05$). The number of years licensed was excluded from the model because it was highly correlated with age. The full model considered was an additive model with the remaining seven demographic variables. The model was reduced by removing the least significant terms until the model was found to be significantly different from the full model. The final reduced model included the sign alternative and age. Table 2 shows the estimates of these variables in comparison with a base value for each variable.

As seen from Table 2, Sign INFO was rated higher than all other alternatives, and Sign i with a circle was rated second highest. The circular border was not found to have had a statistically significant effect on the comprehension rates of Signs i and ?.

Driving Simulator

Subjects were recruited through e-mail lists and posters in community areas located within Corvallis, Oregon and Albany, Oregon. The driving simulator experiment was completed by 42 subjects (28 male and 14 female) with an age range of 21–72 years (average age of 38.7). The highest education level attained by subjects ranged from a high school diploma (2 subjects, or 4.8%) to a Ph.D. (2 subjects, or 4.8%), with the most common education level being some college (16 subjects, or 38.1%). Kilometers driven in the previous year ranged from 0 to 8,047 km (0 to 5,000 mi) (3 subjects, or 7.1%) to more than 24,140 km (20,000 mi) (6 subjects, or 14.3%), with the most common amount of driving being between 16,093 and 24,140 km (10,000 and 15,000 mi) (35.7%). Nine subjects, all of which were female, did not complete the experiment owing to simulator sickness, representing a simulator sickness rate of 17.7%. Of the 42 subjects who completed the experiment, eye-tracking data were not collected for eight subjects, one male and seven females, because of an inability to calibrate the equipment. Thus, 42 subjects provided usable comprehension data and 34 subjects provided usable eye-tracking data.

Table 2. Online Survey Rating Task Reduced Model

Reduced model variable	Level	Estimate	p-value
Sign alternative	**Sign I**	**−7.85**	**0.020**
	Sign i with circle	Base value	—
	Sign INFO	**37.91**	**<0.001**
	Sign ?	**−10.39**	**0.002**
	Sign ? with circle	**−8.63**	**0.011**
Age	**—**	**−0.227**	**0.002**

Note: Bolded values are statistically significant ($p < 0.05$).

Driver Comprehension during Simulated Driving

The 42 usable comprehension responses were scored as 1 if correct, 0.5 if partially correct, and 0 if incorrect, resulting in panel data with a discrete dependent variable. The identical, interrater reliability procedure used for the survey comprehension questions was implemented for the purpose of reducing the driving simulator comprehension data. In addition to the comprehension score, critical confusions were considered.

Multiple recurring, wrong answers were identified in the subject responses to the comprehension questions. The most common incorrect interpretations were that the i symbol indicated that a gas station was nearby or that the blue background indicated it was a hospital sign. Again, the misinterpretation of the sign indicating a hospital nearby was considered a critical confusion. A complete list of the incorrect answers and their frequency is shown in Fig. 9. This table does not include subject responses that failed to include a specific guess.

Three commonalities exist between the different incorrect answers: (1) words that also start with the letter i (interstate or intersection); (2) signs with an identical blue background (hospital or gas station); and (3) signs that have vertical or white symbols in the center (airport or pedestrian).

An ANCOVA test was used to assess differences in the means when considering the factors collected: sign alternative, driver age, level of education, sign order, number of years as a licensed driver, frequency of recreation/pleasure travel, and number of miles driven in the previous year. If significant differences were found, the ANCOVA test was followed by t-tests as recommended by Ramsey and Schafer (2013). A full model was created by including all factors as additive variables. A reduced model was then found

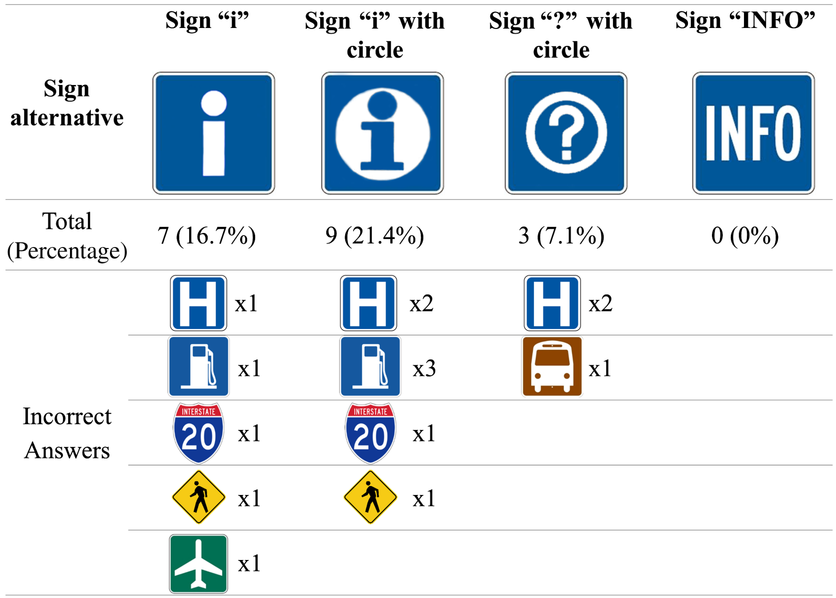


Fig. 9. Incorrect tourist information sign comprehension during simulated driving (FHWA 2009)

Table 3. Driving Simulator Test t-Test p-Values

Sign alternatives	Sign i	Sign i with circle	Sign INFO	Sign ? with circle
Sign I	1.000	—	—	—
Sign i with circle	0.754	1.000	—	—
Sign INFO	**0.021**	**0.032**	1.000	—
Sign ? with circle	**0.042**	**0.024**	**<0.001**	1.000

Note: Bolded values are statistically significant ($p < 0.05$).

by comparing the full model with reduced models until only the significant variables remained. The reduced model that emerged showed a significant impact of the sign type and of the order in which the signs were displayed, $p < 0.001$ and $p = 0.045$, respectively. Two-tail t-tests were performed to determine the comprehension differences between sign alternatives (Table 3).

The statistical results showed that in the driving simulator experiment the "INFO" sign alternative demonstrated the highest comprehension rate. The two Sign i alternatives did not perform differently at a statistically significant level. Sign ? with a circle performed worse than all other alternatives at a statistically significant level. These results, as well as 95% confidence intervals, are shown graphically (Fig. 10).

Visual Attention during Simulated Driving

Differences in visual attention among subjects with correct and incorrect responses were investigated. Initially, descriptive statistics of the total dwell time were calculated for correct (mean 3.96 s and 1.74 s standard deviation), partially correct (mean 3.85 s and 1.30 s standard deviation), and incorrect (mean 4.28 s and 1.88 s standard deviation) subject comprehension. Additionally, a panel linear model was developed to describe the differences between total dwell durations. A model that fit the data was not found for the data, and the score did not have a significant impact on the total dwell time. Ultimately, no connection was found between the visual attention of subjects and the correctness of their responses.

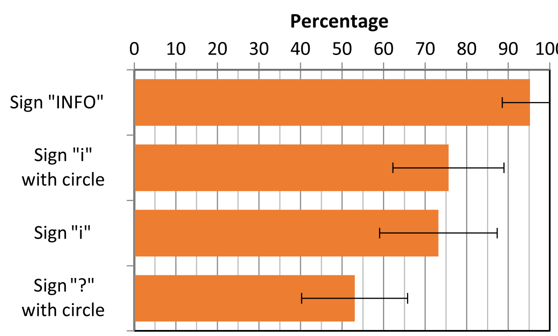


Fig. 10. Driving simulator task averages with confidence intervals

Postdriving Survey Rating Task

The rating task question was identical to the rating task question used in the online survey, with the sole modification that it did not include Sign ? without a circle because that alternative was identified as the least effective alternative in the online survey and was not presented in the driving simulator experiment. Box and whisker plots were created for each of the signs in the rating task (Fig. 11).

In the rating task, Sign INFOperformed best, followed by Sign i alternatives and then Sign ? with a circular border. It was consistently predicted that Sign INFO would have the best comprehension rate compared to the other alternatives. Both random- and fixed-effects models were considered to fit the rating task panel data. A two-way model was chosen to account for the bias that may have occurred owing to subjects' making multiple observations. A Hausman test was conducted on the additive model, and it was found that the random-effects model fit the data better ($p > 0.05$). The number of years licensed was excluded from the model because it was highly correlated with age. The full model, Eq. (3), considered was an additive model with the remaining seven demographics:

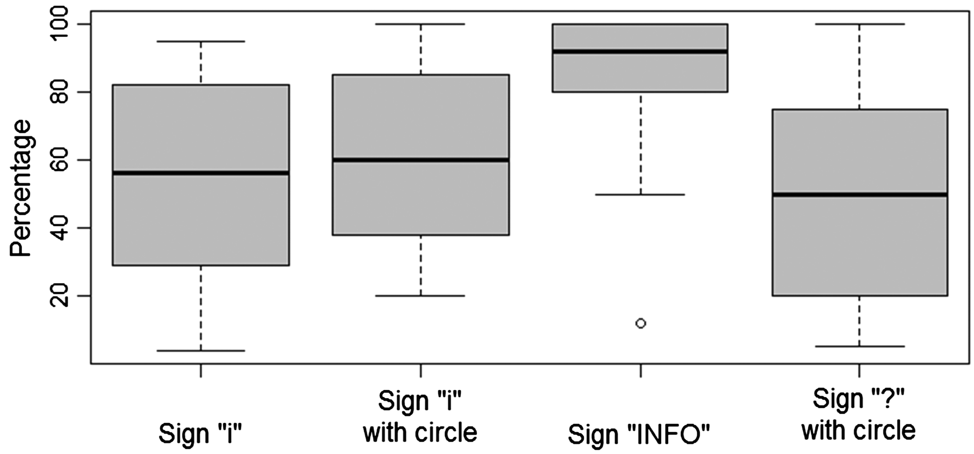


Fig. 11. Sign comprehension scores from rating task

$$\text{Mean Comprehension Score} = \beta_0 + \beta_1 \times \text{Sign Type} + \beta_2 \times \text{Age} + \beta_3 \times \text{Education} + \beta_4 \times \text{Sign Order} + \beta_5 \times \text{Years Licensed} + \beta_6 \times \text{Recreation Travel Frequency} + \beta_7 \times \text{Miles Driven in Previous Year} \quad (3)$$

A reduced model was found by comparing the full model with reduced versions and removing the least significant terms until only significant variables remained. The final reduced model, Eq. (4), included the sign alternative, the miles driven in the previous year, and age:

$$\text{Mean Comprehension Score} = \beta_0 + \beta_1 \times \text{Sign Type} + \beta_2 \times \text{Age} + \beta_3 \times \text{Miles Drivin in Previous Year} \quad (4)$$

Table 4 shows the estimates of these variables in comparison with a base value for each variable.

Sign INFO was rated higher than all other alternatives, and Signs i, with and without a circular border, were rated as second highest (Table 4). Sign ? with a circular border was predicted to have the lowest comprehension rate. The performance difference between Sign INFO, the Sign i alternatives, and Sign ? with a circular border were all significant. There was not a significant difference between the comprehension of Sign i without a circular border and the Sign i alternative with a circular border.

Comparison of Test Methods

Using the research results from this study of alternative information signs in Oregon, the two testing media for traffic sign comprehension (online survey and driving simulator study) were compared. Descriptive statistics for the comprehension questions are compared between the online survey and the driving simulator experiments (Table 5) and for the rating task questions in the online survey and in the postdriving simulator survey (Table 6). The driving simulator comprehension results were considered the baseline for this study because the simulator presents the most authentic simulation of the actual driving task. The results from each testing method follow the same general ranking of sign alternatives (Fig. 12).

Statistical differences were identified between test methods. The Sign i results were significantly different in the rating task ($p = 0.032$), and the Sign ? with circle results were significantly different between the open-ended comprehension methods ($p < 0.001$). Each test found Sign INFO alternative to be statistically superior to all other alternatives. With the exception of the online survey rating task, each test also agreed that the miles driven by the subject in the previous year was the only significant secondary factor. It is also possible that the outliers observed in the online rating task comprehension are an artifact of the interpretation required by the question structure. Because the driving simulator most accurately recreates the driving task, it is the preferred method to test sign comprehension prior to field installation and represents the most authentic experimental task from this study. The open-ended comprehension test also appears to closely match the comprehension task while driving.

Table 4. Driving Simulator Rating Task Reduced Model

Reduced model variable	Level	Estimate	p-value
Sign alternative	Sign i	−8.57	0.112
	Sign i with circle	Base value	—
	Sign INFO	**25.87**	**<0.001**
	Sign ? with circle	**−14.51**	**0.007**
Kilometers driven in previous year	0–8,047 km (0–5,000 mi)	−12.94	0.600
	8,047–16,093 km (5,000–10 mi)	Base value	—
	16,093–24,140 km (10,000–15,000 mi)	−8.43	0.126
	24,140–32,187 km (15,000–20,000 mi)	**−22.09**	**<0.001**
	More than 32,187 km (20,000 mi)	−6.92	0.355
Age	—	**−0.30**	**0.013**

Note: Bolded values are statistically significant ($p < 0.05$).

Table 5. Comprehension Descriptive Statistics from Online Survey and Driving Simulator

	Mean		Standard deviation	
Sign alternative	Online survey	Driving simulator	Online survey	Driving simulator
Sign i	0.75	0.73	0.44	0.44
Sign i with circle	0.84	0.76	0.37	0.43
Sign INFO	1.00	0.95	0.04	0.22
Sign ?	0.73	—	0.44	—
Sign ? with circle	0.79	0.53	0.38	0.40

J. Transp. Eng., 2016, 142(1): 04015036

Table 6. Rating Task Descriptive Statistics from Online Survey and Postdriving Simulator Survey

Sign alternative	Mean: Online survey	Mean: Post driving simulator survey	Standard deviation: Online survey	Standard deviation: Post driving simulator survey
Sign I	42.2	53.7	28.2	27.9
Sign i with circle	50.1	62.3	28.0	26.0
Sign INFO	88.1	87.9	13.3	14.2
Sign ?	39.7	—	27.9	—
Sign ? with circle	41.5	47.8	28.7	29.4

Discussion

This section considers the results in the context of each of the four previously described hypotheses:

1. There is no difference in driver comprehension between each sign alternative;
2. There is no difference in a driver's glance patterns or fixation points between each sign alternative;
3. There is no difference in a driver's glance patterns or fixation points between correct, partially correct, and incorrect responses; and
4. There is no difference in driver comprehension between each sign alternative in the online survey and in the driving simulator.

Following the discussion of the hypotheses, the results are compared to those of previous work.

Hypothesis 1: Traveler Information Sign Comprehension

The primary research objective of this study was to determine the comprehension rates of alternative tourist information signs. Through each of the test methodologies, the Sign INFO alternative outperformed all other alternatives in terms of comprehension at a statistically significant level. The Sign INFO was the only sign alternative considered to meet the ANSI Z535.3 standards of comprehension greater than 85% and critical confusions less than 5%.

Three categories of incorrect comprehension emerged from subject responses. These included a misinterpretation of the blue background, misinterpreting the message communicated by Sign i, and mistaking the i as a different symbol. One of the advantages of word messages is the reduced rates of comprehension errors, whereas symbols can be more easily misunderstood (Katz et al. 2008). From the results of this experiment, 12 of the 15 confusions occurred with Sign i alternatives. This provides evidence that symbol signs developed using single letters might negatively affect comprehension rates.

The general success of the Sign i alternatives may be due to the prevalence of the i symbol in other contexts, in particular, on the Internet and other technologies. The i symbol has been widely adopted on the Internet to inform users of various types of information and, owing to its common appearance, is likely well understood in that context.

Multiple driver-related factors were collected and analyzed to test for differences among subject groups and sign alternatives. The factors analyzed included gender, age, highest level of education completed, number of miles driven last year, frequency of recreation or pleasure travel, primary language, and home state. In all of the tests, except for the online survey rating task, only the miles driven in the previous year variable was found to be significant. In the online survey rating task, the highest level of education completed and the subject's age were statistically significant.

Hypotheses 2 and 3: Traveler Information Sign Glance Patterns

Another consideration was to evaluate the importance of the subjects' dwell times on sign alternatives. The total dwell time was compared in two instances, first, between sign alternatives (Hypothesis 2), and second, between correct, partially correct, and

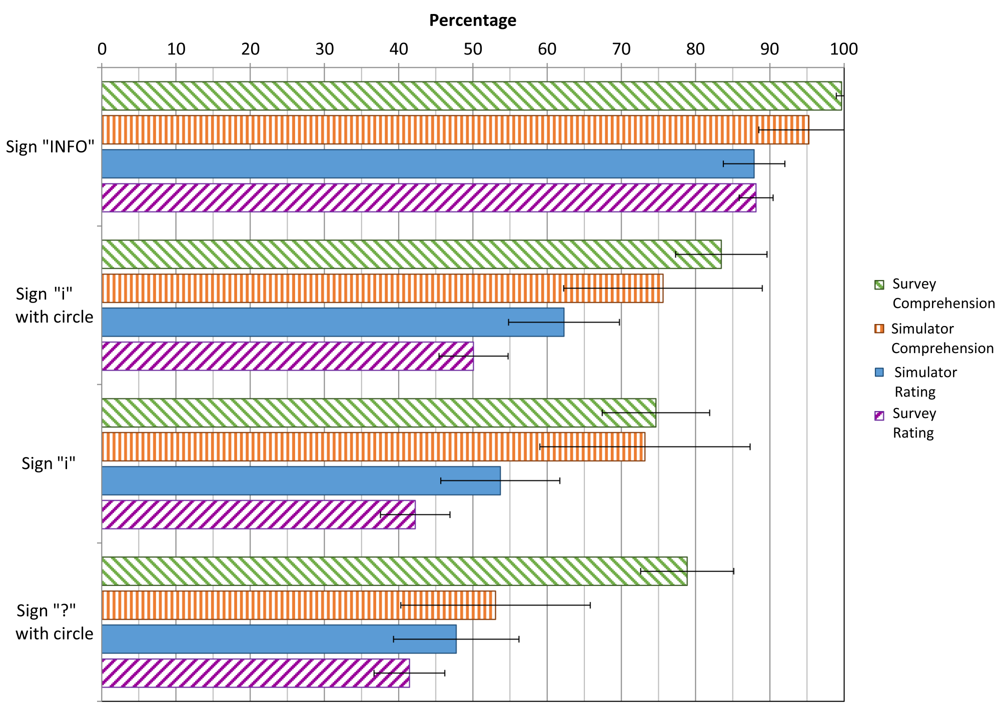


Fig. 12. Testing method results comparison

incorrect responses (Hypothesis 3). The dwell time of subjects was not significantly different among any of the sign alternatives and was not significantly different among correct, partially correct, and incorrect responses. Therefore, subjects did not spend additional time looking at a sign when they did not know the sign's meaning. This could be because the message on each of the alternatives was simple enough that the subject did not find it beneficial to look at the sign more than the time necessary to determine the intended message.

The results from the driving simulator ANCOVA test revealed no significant difference in the glance patterns among sign alternatives, implying that the time required by subjects to read and interpret the "INFO" sign alternative was not significantly different than the time required by subjects to interpret the symbolic sign alternatives. These results fall in line with Ells and Dewar (1979), who found no significant difference in reaction time between symbolic signs and word signs with simple messages like HILL or BUMP.

Hypothesis 4: Experimental Design Effect on Comprehension

Each of the test methods used in this study can also be compared to evaluate the effect of testing procedures on the obtained results. The fourth research hypothesis, which states that there is no difference in driver comprehension between each sign alternative in the online survey and in the driving simulator, was not rejected. There were significant differences in the results of the rating tasks in the online survey and the follow-up survey administered after the driving simulator experiment for both Sign i alternatives and significant differences in the open-ended comprehension tasks in the online survey and simulator for Sign ? with a circular border. This indicates that the testing methodology can affect the results and a more realistic approach is preferred.

Comparison with Previous Sign i Research

The tourist information sign was previously studied by Katz et al. (2008). However, there are several notable differences between the testing procedures implemented in the previous study and this study. The Katz et al. study examined the comprehension (with multiple-choice and open-ended questions) and legibility (with a sign simulator that displays images of signs on a 2D projector) of a multitude of traffic sign alternatives, one of which was the information sign. They specifically considered three symbols, Signs i, ?, and INFO. In contrast, the present study exclusively examined the comprehension of Sign i with open-ended and rating task questions in an online survey and open-ended questions during a simulated driving task. This study specifically considered five symbols: i with and without an inscribed circle, ? with and without an inscribed circle, and INFO.

Katz et al. found that 56% of drivers correctly understood the i symbol, 68% of drivers understood the ? symbol, and 96% of drivers understood the INFO message compared to comprehension rates of 74.7% (Sign i), 72.9% (Sign ?), and 99.7% (Sign INFO) determined in this research. The results for the ? symbol and the INFO message were slightly higher in this research than those reported by Katz et al. (2008). However, the percentage of drivers that correctly comprehended the i symbol was significantly different in the two experiments, which suggests, as one possibility, that the use of the i symbol has increased, leading to increased rates of comprehension. These results may also suggest regional differences in sign comprehension.

It is important to consider that a limitation of Sign INFO is that its legibility degrades at greater distances (Katz et al. 2008). Katz et al. suggests that by increasing the text size to 20.32 cm (8 in.). Sign INFO would be comparable to the legibility distance experimentally determined by Signs i and ?.

Conclusions

This study compared a variety of online survey questions and driving simulation tasks to assess the understandability of alternative tourist information signs in Oregon. In all of these tests, Sign INFO was shown to be the most understandable of the alternatives evaluated in this study by a significant margin, supporting its current usage in the MUTCD over Sign ?. Though, at this time, Sign INFO performed better than alternative forms (e.g., Sign i), the prolific usage of i signs in multiple contexts both domestically and internationally will likely improve its understandability within the context of driving in the future. For example, this study has shown a significant increase (20% in 6 years) in the understandability of Sign i compared to relatively recent findings by Katz et al. (2008). Hence, continued research can evaluate this trend to determine whether Sign i will be a more suitable alternative in the future. The comprehension results showed that Sign i with an inscribed circle outperformed the simple Sign i; hence, future studies could evaluate alternative presentations of i. It is also recommended that public agencies work toward consistency in deploying tourist information signs, which will improve understandability.

While the research presented in this paper is focused on information signs, it has broader relevance for future traffic sign evaluations. First, this study compared four testing modalities and showed those modalities to be relatively valid (the resultant rank ordering of each sign alternative based on comprehension rates are similar, but the magnitude of the comprehension rates vary between modality). Notably, however, this work showed that the comprehension rates observed in the driving simulator during a simulated driving task are similar but consistently lower than those observed in open-ended questions administered in an online survey. The study also showed that the same symbol, when presented differently, can have improved understandability. Finally, this work showed that the understandability of a sign may not be permanent, with a significant shift possible in as little as 6 years. For certain types of signs it may be useful to periodically reevaluate comprehension rates.

Acknowledgments

This project was jointly funded by Travel Oregon and the Pacific Northwest Transportation Consortium (PacTrans). The authors would like to recognize the contributions of time and technical expertise provided by Scott West and Carole Astley of Travel Oregon, Natalie Inouye of Travel Lane County, and Michael Kimlinger with the Oregon Department of Transportation (ODOT). This study was conducted with support from the Oregon State Center for Healthy Aging Research, Life Registry.

References

ANSI (American National Standards Institute). (2011). "ANSI Z535.3 criteria for safety symbols." National Electrical Manufacturers Association, Rosslyn, VA.

Babbitt Kline, T. J., Ghali, L. M., Kline, D. W., and Brown, S. (1990). "Visibility distance of highway signs among young, middle-aged, and older observers: Icons are better than text." *Hum. Factors: J. Hum. Factors Ergon. Soc.*, 32(5), 609–619.

Benjamini, Y., and Yekutieli, D. (2001). "The control of the false discovery rate in multiple testing under dependency." *Ann. Stat.*, 29, 1165–1188.

Castro, C., Horberry, T., and Gale, A. (2004). "The effect of semantic and repetition priming on traffic sign recognition." *Vision in vehicles*, Univ. of Derby, Derby, U.K.

Dean Runyan Associates. (2014). "Oregon travel impacts 1991–2013p. Statewide preliminary estimates." Oregon Tourism Commission, Salem, OR, 1–239.

Dewar, R. (1988). "Criteria for the design and evaluation of traffic sign symbols." *Transportation Research Record 1160*, Transportation Research Board, Washington, DC, 1–6.

Ells, J. G., and Dewar, R. E. (1979). "Rapid comprehension of verbal and symbolic traffic sign messages." *Hum. Factors: J. Hum. Factors Soc.*, 21(2), 161–168.

FHWA. (2000). "Manual on uniform traffic control devices." U.S. DOT, Washington, DC.

FHWA. (2003). "Manual on uniform traffic control devices." U.S. DOT, Washington, DC.

FHWA. (2009). "Manual on uniform traffic control devices." U.S. DOT, Washington, DC.

Hurwitz, D., Monsere, C., Marnell, P., and Paulsen, K. (2014). "Three- or four-section displays for permissive left turns? Some evidence from a simulator-based analysis of driver performance." *Transportation Research Record 2463*, Transportation Research Board, Washington, DC.

Katz, B. J., Dagnal, E. E., and O'Donnell, C. C. (2012). "Evaluation of comprehension and legibility of international and domestic nonstandard symbol signs." *Transportation Research Record 2298*, Transportation Research Board, Washington, DC, 46–60.

Katz, B. J., Hawkins, G., Kennedy, J. F., and Howard, H. R. (2008). "Design and evaluation of selected symbol signs." *TPF-5(065)*, Federal Highway Administration, McLean, VA.

Konstantopoulous, P., Chapman, P., and Crundall, D. (2010). "Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving." *Accid. Anal. Prev.*, 42(3), 827–834.

Marnell, P., Tuss, H., Hurwitz, D., Paulsen, K., and Monsere, C. (2013). "Permissive left-turn behavior at the flashing yellow arrow in the presence of pedestrians." *Proc., 7th Int. Driving Assessment Conf. Compendium*, Univ. of Iowa, Ames, IA, 488–494.

Paniati, J. F. (1988). "Legibility and comprehension of traffic sign symbols." *Proc., Human Factors and Ergonomics Society 32nd Annual Meeting*, Sage, Thousand Oaks, CA, 568–572.

Ramsey, F. L., and Schafer, D. W. (2013). *The statistical sleuth: A course in methods of data analysis*, Brooks/Cole, Boston.

Smiley, A., MacGregor, C., Dewar, R. E., and Blamey, C. (1998). "Evaluation of prototype highway tourist signs for Ontario." *Transportation Research Record 1628*, Transportation Research Board, Washington, DC, 34–40.

UNWTO (United Nations World Tourism Organization). (2000). "Recommendation on the tourist information sign." *Executive Council Decision CE/DEC/6*, New York.

Walker, R. E., Nicolay, R. C., and Stearns, C. R. (1965). "Comparative accuracy of recognizing American and international road signs." *J. Appl. Psychol.*, 49(5), 322–325.

Wolff, J. S., and Wogalter, M. S. (1998). "Comprehension of pictorial symbols: Effects of context and test method." *Hum. Factors: J. Hum. Factors Ergon. Soc.*, 40(2), 173–186.

Zwaga, H. J. (1989). "Comprehensibility estimates of public information symbols; their validity and use." *Proc., Human Factors and Ergonomics Society Annual Meeting*, Vol. 33, Sage, Thousand Oaks, CA, 979–983.