Title
Evaluate the Understandability of Information Display Board Signs Using a Driving Simulator Experiment

Journal
International Journal of Human-Computer Interaction, 38(17)

ISSN
1044-7318

Authors
Wang, Pei
Zhang, Tingting
Zhou, Xiao
et al.

Publication Date
2022-10-21

DOI
10.1080/10447318.2021.2014180

Peer reviewed

Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

Evaluate the Understandability of Information Display Board Signs Using a Driving Simulator Experiment

Pei Wang, Tingting Zhang, Xiao Zhou, Sanaz Motamedi & Ching-Yao Chan

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2021.2014180

Evaluate the Understandability of Information Display Board Signs Using a Driving Simulator Experiment

Pei Wang[a], Tingting Zhang[a], Xiao Zhou[b], Sanaz Motamedi[a,c], and Ching-Yao Chan[a]

[a]California PATH, UC Berkeley, Richmond, CA, USA; [b]School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan, China; [c]Industrial and System Engineering, University of Florida, Gainesville, FL, USA

ABSTRACT
The main research question of this study was how to apply the design guidelines for traditional Changeable Message Signs (CMSs) to the design of the full-color and full-matrix LED Information Display Boards (IDBs). Three categories of messages were evaluated in this study, including (1) travel-time messages, (2) transit travel-time messages, and (3) Graphic Route Information Panels (GRIPs). A driving simulation program was developed based on real-world videos and used to evaluate the designed signs in terms of understandability and helpfulness for decision-making. A total of twenty-four local commuters participated in the driving simulator experiments. Results show that: (1) the perceived easiness of five-line travel-time message is significantly lower than the baseline three-line message, (2) the perceived helpfulness of the transit logo is significantly higher than the generic symbol, and (3) there is no significant difference between the drivers' comprehension of the three-line message and the single-link GRIPs.

1. Introduction

To better communicate travel time information and reduce accidents and congestions, the California Department of Transportation (Caltrans) has installed changeable message signs (CMS) with full-color and full-matrix LED display technology along the Interstate 80 (I-80) corridor. These CMSs are also referred to as Information Display Boards (IDBs). The IDBs were installed to display traveler information messages such as congestion level and expected travel time to enhance drivers' knowledge of the downstream traffic conditions and provide information about alternative routes and alternative transportation modes (i.e., driving vs. public transit). An IDB in the field is shown in Figure 1. The visible optical area is 157 1/2 by 196 7/8 inches. The resolution is 216 by 270 pixels. Compared with the traditional LED technology-based CMSs, which have a pixel spacing of 2.75 inches, the IDBs have a much smaller pixel spacing of 0.73 inches. The higher resolution and smaller pixel spacing of the IDBs allow the flexibility of displaying more complex graphics.

It would require a message that exceeds the two-phase, four-unit limit as recommended in the MUTCD to convey the congestion downstream and the location, length, severity of the congestion, and recommended alternative route. Here raises the question of whether the design of IDBs is subject to the guidelines for the traditional CMSs. More specifically, in the current study, we are exploring three research questions about the IDB design. (1) Whether more than three units of information with two virtual phases of messages could be displayed on one sign? (2) Whether the text-symbol design outperforms the text-only design for the transit travel-time message? (3) How well graphic information panels (GRIPs) perform to communicate the traffic congestion information through the IDBs?

1.1. Changeable message sign (CMS) and information display board (IDB)

According to the definition by the Federal Highway Administration (FHWA) Manual on Uniform Traffic Control Devices (MUTCD) (FHWA, 2009), a changeable message sign (CMS) is a traffic control device that is capable of displaying one or more alternative messages. CMSs are an essential part of the driver information system and an important link between transportation agencies and the driving public. They can be used to effectively manage travel, identify current and anticipated roadway conditions, and regulate access (National Research Council (US), National Cooperative Highway Research Program, American Association of State Highway and Transportation Officials & United States, 2012). The California MUTCD 2014 edition, which is in conformance with the national MUTCD 2009 edition, guides the letter height, character spacing, resolution, and other characteristics of traditional CMSs (California Department of Transportation, 2014). Each message shall consist of no more than two phases. A phase shall consist of no more than three lines of text. The message should be legible from a minimum distance of 600 feet for nighttime conditions and 800 feet for normal daylight conditions. According to the California MUTCD, with a full matrix type of technology, the CMS can duplicate a standard

CONTACT Xiao Zhou zhouxiao@whut.edu.cn School of Mechanical and Electronic Engineering, Wuhan University of Technology, 122 LuoShi Road, HongShan District, Wuhan, Hubei 430070, China

Figure 1. IDB in the field.

sign using standard symbols, standard alphabets, letter forms, route shields, and other typical sign legend elements. The use of the full-matrix type of technology is encouraged for increased legibility and enhanced recognition.

1.2. CMS configurations and impacts on drivers

1.2.1. Message length

As there is limited time for drivers to read a sign, controlling the message length is critical to the correct message comprehension. Message length is about the number of words and the number of information units included. Dudek (2004) stated that the appropriate message length is affected by (1) the amount of time the driver has in the legibility zone of the CMS; and (2) the amount of activity in the traffic stream that the driver must attend to, such as reading signs and lane positioning. Drivers can read and comprehend a well-designed CMS message when displayed at a rate of 2 seconds per unit of information (Dudek & Ullman, 2006; Dudek et al., 2007). Campbell et al. (1998) reported that character resolution could affect the readability of the texts on a CMS. For characters smaller than approximately 22 arcminutes, a 7×9 matrix resolution led to shorter reading times and fewer reading errors than a 5×7 matrix resolution.

Jamson et al. (2005) studied the effects of configurations of bilingual CMSs on driver behavior, and safety. Results indicated that the drivers could read one and two-line monolingual signs and two-line bilingual signs without changing their driving behavior. However, the drivers significantly reduced their speed to read four-line monolingual and four-line bilingual signs, accompanied by an increase in the headway to the vehicle in front.

Lai (2010) investigated the effects of the number of message lines (i.e., single, double, and triple) of Chinese CMSs on participants' response performance through a laboratory experiment. The analysis results showed that the number of message lines was a significant factor in the response time to CMSs. Participants took less response time for the double-line message than for the single and triple-line message. However, it was recommended not to overload drivers with too much information. Xu et al. (2020) used dynamic simulation experiments to assess the influence of information volume (i.e., number of roads displayed) on graphical CMSs. The results indicate that the drivers' comprehension accuracy decreased significantly when the number of roads increased to six. It is recommended that five is the maximum number of roads to be shown on a graphical CMS.

One very important application of CMSs is traffic incident management, for which the messages cannot be presented in a single phase and need to be presented in successive phases. Dutta et al. (2004) analyzed the factors that affected the readability and comprehension of multiphase messages presented on CMSs during a simulated driving task and identified the factors that maximize driver performance. The study underscores the need to reduce the phase exposure time as long as the drivers do not fail to understand the messages and suggests that the biphasic messages should be repeated, if possible, while the driver is in the legibility zone.

1.2.2. Use of graphics

Ells and Dewar (1979) conducted two experiments to measure the time required to comprehend the traffic signs. The results indicated that the signs with symbolic messages could be understood more quickly than those with verbal messages. Lai (2010) investigated the effect of color schemes (one, two, and three) of Chinese CMSs on participants' response performance. Results showed that the participants responded faster for two-color than for one and three-color schemes. Yan and Wu (2014) evaluated the effectiveness of CMS on changing drivers' decisions based on a driving simulation experiment. It was found that the drivers are more willing to change routes with the CMS with graphics than in the text-only format.

Besides, the text-only CMS raised responding time and increased abrupt deceleration behaviors during the lane changings. Zhao et al. (2019) surveyed a thousand private and taxi drivers in China about their preferred CMS contents and formats. The results indicated that, in normal weather conditions, drivers with working purposes mostly preferred to receive the information on a congested traffic condition in the text-only format. Drivers preferred to receive the route diversion suggestions in a graph-only form in foggy weather and the qualitative delay time in a text-graph form in normal weather conditions. Roca, Tejero, et al. (2018) assessed the difficulties of adults with dyslexia acquiring the information shown in CMS and provided evidence to discuss the use of pictograms as potential countermeasures. The results revealed that adults with dyslexia, despite possible compensation effects, find difficulties reading text messages in CMS (shorter legibility distances, longer reading time, and increased cognitive effort). In contrast, there were no such differences in the recognition of pictograms. Therefore, pictograms proved to be more resistant to individuals' adverse conditions, such as dyslexia, compared to text messages. P. Wang et al. (2019) conducted a lab test to evaluate the understandability of different designs of IDB signs (e.g., with symbol vs. without symbol). The results showed that the designs with both symbols and texts had higher understanding accuracy and required shorter viewing time than their text-only counterparts.

Ng and Chan (2007) pointed out that the influence of a symbol's visual features (e.g., color, shape, size) and cognitive features (familiarity, concreteness, complexity, meaningfulness, semantic distance) should be considered while using symbols. Similarly, Ben-Bassat and Shinar (2006) evaluated the influence of ergonomic principles of familiarity, standardization, and symbol-concept compatibility on traffic sign comprehension. The results showed that understanding of the evaluated symbol concepts was highly correlated with familiarity.

Roca, Insa, et al. (2018) used a driving simulator to present either single-word or pictorial messages on CMSs to a group of drivers. The reported results, to some extent, were inconsistent with the previous researches with the findings that pictorial information had advantages over text information in traffic signs. Single-word messages on CMSs could be read significantly farther and were identified with higher accuracy than pictograms. Besides, drivers approaching CMSs with text messages exhibited lower speed variability and dedicated fewer glances and less glancing time at the sign than pictogram messages. In other words, the text messages were associated with higher legibility distance and understand accuracy and demanded fewer cognitive resources, allowing more stable control over the vehicle speed and requiring fewer visual demands.

1.2.3. GRIPs

Alluri et al. (2017) conducted a focus-group study to evaluate different express lane guide signs. It was concluded that additional real-time information on travel time and average speed on express lanes helped drivers to decide whether they wanted to use the express lanes. Compared with traditional CMS, Graphic Route Information Panels (GRIPs) have the potential to display greater details of traffic information. GRIPs use a combination of text, colors, and a map representing one or more roadways to convey the location and severity of the congestion. The research on GRIP started in Japan (Takeda & Kujirai, 1999) and then became popular in Europe (Richards et al., 2004) and China (Gan et al., 2008). Presently, there is no information available regarding GRIP in the MUTCD.

Crundall et al. (2011) addressed two questions concerning the presentation of GRIPs using an eye-tracking approach. One question was whether the map should be represented in an allocentric or egocentric form. Allocentric maps were presented with a north-up orientation, while egocentric maps were oriented such that the entry point was always in the middle of the bottom (a heading-up orientation). A second question was how closely the map should represent the real world, which could be viewed with a continuum ranging from topographic to schematic. The topographic approach tries to emulate the real world, with roads and junctions representing their true spatial location. The schematic approach removes unnecessary detail and displays only task-relevant information. Results of behavioral measures favored the schematic-egocentric orientation. In addition, schematic maps received more accurate responses than topographic maps, and egocentric maps were responded to more accurately than allocentric maps. The eye movement measures also supported the benefits of the schematic-egocentric maps, with total dwell time (TDT) being greater on topographic and allocentric trials. The authors also recommended that the designers of GRIPs need to be aware of the problems of an extremely long path, which will encourage needless line tracing and potentially result in longer off-road gazes. The influence of orientation on processing time was also studied in both the Netherlands and Germany. The research findings recommended the use of a bottom-to-top orientation for GRIPs (Alkim & Schenk, 2001; Schonfeld et al., 2000).

Lai (2012) concluded that GRIP with road color was the easiest to understand, and the GRIP with both road color and travel-time can be used for a simple road network. Additionally, regarding colors on GRIPs, some guidelines were developed, such as consistency in the use of color codes and avoidance of using colors from extreme ends of the color spectrum (Brockmann, 1991). In a study conducted by Texas Transportation Institute (Ullman et al., 2009), three colors were used to represent congestion levels on the GRIP: green (normal operating speed), yellow (slow traffic), and red (stop-and-go). The three colors have been well interpreted by drivers and help drivers decide on selecting alternative routes. Alkim and Schenk (2001) conducted a simulator experiment to compare the comprehension of regular CMS and GRIPs. It was found that increased message complexity of the GRIPs could reduce driving speed.

1.2.4. Influence of drivers' difference

Diop et al. (2019) adopted an extended Technology Acceptance Model (TAM) to predict and explain road users' intention to use CMS information. The results indicated that information quality directly affected perceived usefulness, perceived ease of use, and attitude toward route diversion as suggested on the CMS. Drivers' familiarity with the roadway network also had a positive effect on their attitude toward route diversion. Drivers' gender and native language have been suggested to potentially account for comprehension variability of traffic signs (Al-Madani & Al-Janahi, 2002; Ou & Liu, 2012). The study by Ou and Liu (2012) showed that Taiwanese had a significantly better understanding of traffic signs compared to Vietnamese, which are the second-largest ethnic group in Taiwan. Ng and Chan (2007) studied the effect of driving experience and concluded that years of active driving and hours of driving in the past 12 months were not related to drivers' comprehension level. Gan and Ye (2013) explored urban freeway diversion responses to a type of CMS that explicitly provided travel times for both freeways and local streets in China through an on-site questionnaire survey. It was found that drivers' years of driving experience serve as a positive factor in their diversion responses. The study conducted by Dewar et al. (2001) found that older drivers understood traffic signs more poorly than younger drivers. However, participants' age was found to have no significant influence on the signs' understandability, according to the results of the IDB lab-testing conducted by P. Wang et al. (2019).

1.3. Methods for evaluating understandability of traffic signs

Dewar (1990) examined several criteria for traffic sign symbols for the design and evaluation process s from a survey of traffic sign experts in different countries. Understandability was rated as the most important factor, with conspicuity the second. Reaction time, legibility distance, and glance legibility were rated as equally important. In a study conducted by Mackett-Stout and Dewar (1981), glance legibility, legibility distance, comprehension, and preference were used to identify the adequacy of signs. Significant positive correlations were found among the first three measures. Using a driving simulator, Charlton (2006) compared multiple measures of sign processing, including attentional conspicuity, search conspicuity, implicit and explicit recognition, dynamic comprehension, static comprehension, and sign priming. It was found that attentional conspicuity, search conspicuity, and static comprehension were the most reliable indicators of a sign's overall performance. Signs are considered acceptable when at least 67% understanding accuracy is achieved, according to ISO standard (ISO, 2011). In the US, the threshold is 85%, according to the American National Standard Institute (ANSI). From existing literature, it could be concluded that understandability/comprehension is one of the most important measures for evaluating traffic signs.

Various methods have been applied in the studies to understand drivers' perception of freeway signs (Alluri et al., 2017; Gan & Ye, 2013; Macdonald & Hoffmann, 1991; Ng & Chan, 2007; Shinar & Vogelzang, 2013). Common methods include capturing drivers' eye movement and visual attention, using verbal reports while driving, asking drivers to recall a sign after passing it, and recording drivers' driving behavior. Each method has its specific advantages and limitations (Martens, 2000). The main advantage of the eye movement method is that it is relatively free from bias under most circumstances despite the instructions. The main disadvantage is that despite fixation on a certain object, attention can be directed to another location (Gao et al., 2006). Using verbal reports while driving can provide insights into the driver's internal process, complementing the information from eye movement studies (Charlton, 2006). However, the verbal-report method has been mostly used to evaluate a sign's conspicuity rather than its understandability. In the practice of recalling after passing a sign, drivers are asked whether they perceive certain sign information, with which the experiment should be set up in a way that the participants have a clear picture of what to report. In addition, the reports should be requested immediately after the trial to keep the factor of "forgetting" to a minimum level. Fisher argued that the true measure of a sign's effectiveness was not the recall or recognition score but the extent to which the sign content would affect drivers' preparedness for and subsequent responsiveness to events (Fisher, 1992).

1.4. Objectives of this study

In this study, the main research question is how to apply the design guidelines for traditional CMSs to the design of IDBs. More specifically, the research team was interested in determining the impact of message length (e.g., five-line and six-line vs. three-line), the use of graphics, and displaying GRIPs on the IDBs. Therefore, three categories of messages were evaluated. The first category is travel-time messages with two virtual phases. The second category is transit travel-time message. The third category is Graphic Route Information Panel (GRIP), which presents color-coded congestion levels and travel time. It is essential to design the IDBs that communicate unambiguous messages. In this study, we focus on exploring the understandability of different IDB designs. We evaluate how easy it is for drivers to understand each design and how helpful it is for drivers to make decisions regarding route selection. This study was carried out with a driving simulator. We collected the feedback and responses of test subjects in a simulated environment representing real-world driving conditions. The experiment combined several methods, including eye movements, recall after passing the sign, and drivers' preparedness for arriving at the destination of each trip. It is hoped that the findings of this study shed some light on drivers' comprehension and preferences on designs of CMSs with advanced LED display technologies and could be further used as design guidelines for both researchers and traffic engineering practitioners.

2. Methods

2.1. Tested messages

A total of 16 messages in three categories were designed and tested in this study. For each message, the font size height was 24 pixels, the font size width was 14 pixels, and the character spacing was 4 pixels. Colors for legends and background were following the guidelines on CA MUTCD (California Department of Transportation, 2014). For most messages, a black background was used, and the color of the legend was yellow. For a few messages, a green background was used, and the color of the legend was white.

2.1.1. Up-to-six-lines travel-time messages

As shown in Figure 2, one three-line, one five-line, and one six-line message were included. The three-line message the MUTCD recommends serves as the baseline for comparison purposes. The five-line message has two virtual phases with one horizontal line in between, as shown in Figure 2b. The six-line message also has two virtual phases with traveling information about one destination in each phase, as shown in Figure 2c.

2.1.2. Transit travel-time messages

A transit travel-time message conveys travel time from a certain starting point to various destinations by taking public transportations. The reason for providing transit travel-time on a freeway is to inform drivers that taking public transit could be faster than driving at certain times (e.g., peak hours). So that drivers may consider taking alternative transportation modes other than driving. Both bus (i.e., AC Transit) and subway (i.e., BART) related messages were evaluated in this study.

2.1.2.1. AC Transit. AC Transit is a bus service operating in the east bay of the San Francisco Bay Area. There were three alternative graphical designs for the AC transit travel-time messages, as shown in Figure 3. The first design (Figure 3a) has the AC Transit logo on top of the display. Below the logo, it has texts of the starting point and the destinations with a horizontal line in between. In the second design (Figure 3b), an additional bus symbol is on the left side of the AC Transit logo. The bus logo is intended to help drivers who are unfamiliar with AC Transit to understand this message better. In

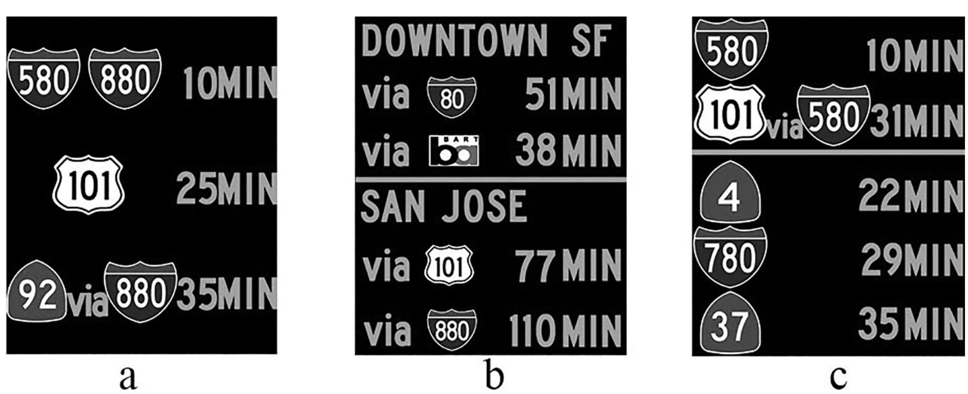


Figure 2. Up to six lines travel-time messages.

Figure 3. AC Transit travel-time messages.

the third design (Figure 3c), there are only “AC TRANSIT” texts without any symbols.

2.1.2.2. BART. BART is the Bay Area Rapid Transit. Two types of messages about the BART service were evaluated, as shown in Figure 4. One is the BART travel-time message, which has two alternative graphical designs. The first design (Figure 4a) has the BART logo on top of the display. Below the logo, it has texts of the starting point and two destinations, with a horizontal line in between. In the other design (Figure 4b), there are only texts of “BART” without the logo. The other type is the BART special message, conveying the information about abnormal situations of BART service, which also has two alternative graphical designs. The first design (Figure 4c) uses the light rail transit station symbol (I-12 on CAMUTCD) (California Department of Transportation, 2014), indicating BART, with texts blow the symbol expressing no service. The other design (Figure 4d) uses the BART logo with similar texts below the logo.

2.1.3. Graphic route information panels (GRIPs)

The GRIPs display traffic conditions of interchanges or important destinations in the schematic format, as recommended by (Crundall et al., 2011), which are limited to three interchanges or destinations. Both single-link GRIPs and dual-link GRIPs are included in this study.

2.1.3.1. Single-link GRIPs. According to (J. H. Wang et al., 2006), specific information on CMS is preferred. We added a road-work legend to indicate the causes of the congestion. Therefore, two design factors were considered for the single-link GRIPs: (1) orientation of approaching destinations, bottom-top or top-bottom, and (2) with or without the road-work legend in conjunction with the color-coded congestion levels. As shown in Figure 5, all single-link GRIPs have either up-arrow or down-arrow. For all single-link GRIPs, it assumes driving on eastbound of I-80 corridor, and Crockett is the east-most destination on the sign.

2.1.3.2. Dual-link GRIPs. As shown in Figure 6, one design factor was considered for dual-link GRIPs: the position of the destinations: one design (Figure 6a) has the destinations in the middle of the two links, and the other design (Figure 6b) has the destinations on the left side. The orientation of the two dual-link GRIPs is up-arrows.

2.2. Participants

Participants were recruited by sending group e-mails to the UC Berkeley communities or distributing flyers in nearby shopping malls along the I-80 corridor. All participants were required to hold a valid California driver’s license. Participants were compensated US$20 per hour for their participation in the study, which lasted for about 1.5 hours. Due to the high percentage of native Spanish-speaking residents along the corridor, we investigated the effects of drivers’ native language (i.e., English vs. Spanish) on the understandability of IDB signs. Both native Spanish-speaking participants and native English-speaking participants were recruited.

Before implementing this driving simulator experiment, the research team conducted a static lab testing to investigate drivers’ understanding of the IDB signs (P. Wang et al., 2019). In the lab testing, 48 local commuters participated. Based on the results of the lab testing, we revised the design of all signs. We also found that age had no significant effect on drivers’ understanding of the IDB signs. Therefore, age was not included as an independent variable in this follow-up driving simulator experiment. In total, twenty-four participants took

Figure 4. BART travel-time messages.

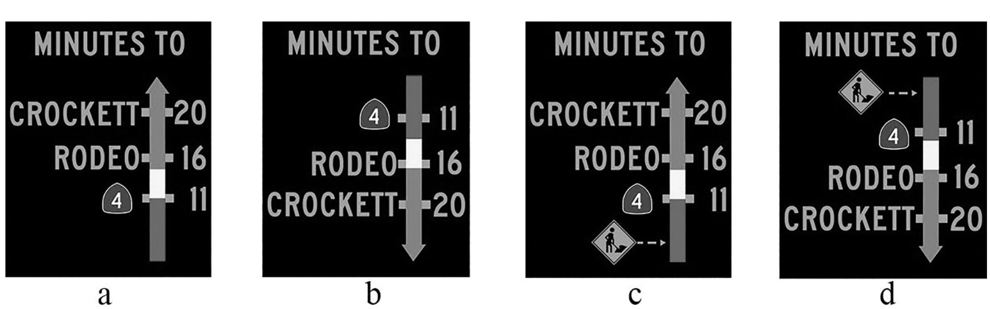


Figure 5. Single-link GRIPs.

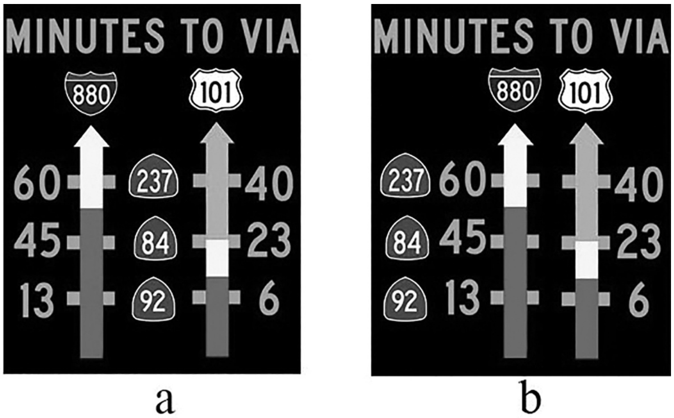


Figure 6. Dual-link GRIPs.

part in the study, including 12 native English speakers and 12 native Spanish speakers. The mean age of all participants was 47.71 years old (SD = 10.35). The gender of participants was balanced, with 12 males and 12 females. The average of their driving experience was 23.48 years (SD = 10.62). All participants were local commuters, each of whom averagely drove on the I-80 corridor 24 times per month (SD = 18.4). The informed consent form was obtained from each participant.

2.3. Driving simulator

This study used a driving simulator located at the University of California, Berkeley. As shown in Figure 7a, the simulator model is Force Dynamic 401CR, mounted on a floor, allowing for continuous movement in four axes. The simulator is equipped with a triple-monitor display and surround sound. The experiments require participants to make physical and attentional efforts on the driving tasks, making the simulation more representative of driving a car in reality. A Tobii 4 C eye tracker was mounted on the middle monitor's bottom frame to capture drivers' eye movement data.

2.4. Driving simulation program

A driving simulation program was developed to display the designed signs in the recorded real-world videos. Original video and vehicle motion data were recorded by driving an instrumented vehicle on the I-80 corridor during the daytime in normal and sunny weather conditions. Three Logitech C922X cameras were mounted to record the views in the front, left-front, and right-front direction of the vehicle to generate a seamless field of view of 150 degrees. The resolution of each video was 1920 × 1080. The frequency was 30 frames per second. The vehicle motion data were recorded at 100 Hz with a 3DM-GX4-45 IMU sensor. Around two minutes of driving data were recorded for each experiment trial. To avoid repetition, each IDB location was recorded multiple times to test different signs. During recording, the instrumented vehicle was always driven on the far-right lane before approaching the IDBs, so that the viewing angle for each sign would be consistent. The driving speed was maintained around 60mph for the travel-time messages and the transit travel-time messages, while it was around 30mph for GRIPs. The driving speed was set to ensure that participants would have enough time to read each sign in the driving simulation.

Image processing techniques were used to integrate the designed signs onto the recorded videos. Firstly, the edge of the IDB sign was detected at the sub-pixel level and then tracked in subsequent frames using an affine model to recognize the position of the sign inside each video frame, as shown in Figure 7b. Then, the designed sign image was mapped onto each frame and replaced the original sign image using the projective transformation. Other techniques were applied to make the signs' visual effect as realistic as seen on the road.

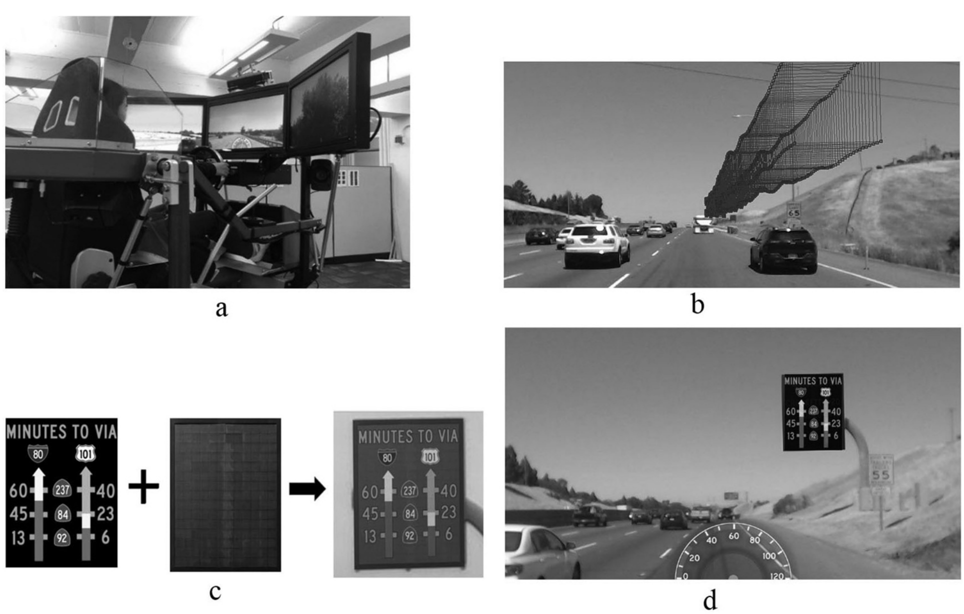


Figure 7. Experiment set-up.

Comprehensive methodological details of the image processing method are reported in another paper.

The processed images and the motion data were assembled into a simulation program to be driven by the participants. The simulation program played the videos on three monitors of the driving simulator and sent motion signals in roll, pitch, and yaw directions to the simulator's motion platform in the same timeline. A filter was added to adjust the vertical and horizontal movement to make the driving experience resemble real-world driving. The input from the participants controlled the speed of playing the video and the movement of the simulator. A speedometer was displayed on the screen to indicate the current driving speed. The driving speed of 55–75 mph was recommended for trials showing travel-time and transit travel-time messages, and the upper limit of speed was 75mph. The 25–45 mph range was recommended for trials showing GRIPs, and the upper limit was 45 mph. To control the viewing time for each sign, an empirical test was conducted to determine a time-point from which drivers with normal vision can see the sign. From this time point, the simulation program started to display the designed sign. Once the program began to show the designed sign, the speed was maintained at a constant value so that the sign could be displayed for a fixed duration of time for each participant. According to the results of viewing time from the lab testing (P. Wang et al., 2019), the fixed duration for the travel-time message and transit travel-time messages was 15 seconds, while it was 20 seconds for GRIPs.

Participants had control over the driving speed through input from the throttle and brake of the simulator. When the throttle was pressed, the video replay speed and the simulator's motion would increase simultaneously. In this way, the high-fidelity simulated driving is achieved, with the simulator's motion similar to a vehicle moving on the real-world freeway geometry. However, they had no control over the steering wheel due to the lane positioning of the vehicle in the prerecorded videos.

2.5. Testing procedure

The testing procedure included (1) screening, (2) informed consent, (3) safety orientation, (4) practice, and (5) main experiment. Details of each procedure are described as follows. After arriving at the simulator room, participants were asked to read and sign an informed consent form. Then participants were asked to complete a screening form. It had twelve statements that evaluated participants' age, driving experience, and physical conditions to determine their qualification for attending the testing. Participants must be over 18 years old, hold a valid California driver's license, and frequently travel on the I-80 corridor to be qualified. For the safe operation of the driving simulator, participants should also meet physical requirements regarding their height (i.e., over 4 feet and under 6.5 feet), weight (i.e., under 250 lbs), history of bone injury (i.e., had not broken any bones in the past year), etc. Participants were asked to initial under each statement. Only the participants who met all the criteria were included in the testing.

After the screening, participants were given a safety orientation regarding how to operate the steering wheel and pedals of the simulator and what to do in case of an emergency. Then participants were led into the driving simulator and comfortably seated with the seatbelt fastened. After being seated, participants were shown a static image of the IDB sign on the middle screen of the simulator and asked whether they could see the content on the sign. The image had the same resolution and design features (e.g., font size, spacing) as the first frame of each designed sign shown in the video programs. This test ensured that all participants could read the content on each designed sign once displayed in the video programs. After that, the experimenter calibrated the Tobii eye tracker. A local map was then used to explain the locations of the IDB signs along the I-80 corridor and the neighboring cities that they would use as destinations during the experiment.

Afterward, it was a practice trial, through which the participants learned how to operate the simulator and get familiar with the testing procedures. At the beginning of the trial, participants were told the destination of this trial (e.g., downtown Oakland) and asked to pay attention to an electronic sign shown on the right-hand side of the freeway. Then they were asked to press the throttle to start driving. Participants' input to the throttle and brake controlled the video's playback speed, therefore the traffic speed and motion of the driving simulator. They were instructed to maintain the driving speed within a specific range between 55 mph and 75 mph for the practice trial. After completely passing the sign, the simulation program of the practice trial ended. Then participants were asked questions about (1) general information of the message; (2) specific information associated with the destination; (3) helpfulness of symbols, if any, on a scale of 1 to 5, 1 being not helpful and 5 very helpful; and (4) easiness to understand the sign, on a scale of 1 to 5, 1 being very hard and 5 very easy.

After the practice session, it was the main experiment, in which each participant evaluated 16 displays, including all the messages described in the section of "Tested Messages." The evaluation procedure for each message was the same as the practice trial. The sequence of all trials was randomized for each participant. To ensure participants' attention and fatigue levels were in a relatively reasonable range throughout the experiment, participants were asked to get off the simulator and take a break after the evaluation of every 4 signs. Meanwhile, to avoid deterioration in the accuracy of the eye tracker, re-calibration procedure was implemented after the break. It took around 3 to 4 minutes to complete each trial. The whole experiment took about 90 minutes.

2.6. Experimental design

Besides the sign design, drivers' gender and native language were also controlled in the experimental design. As every participant evaluated all the signs, the experiment was a repeated measure design. It considered design factors as the main independent variables. The summary of design factors for each category of the messages is shown in Table 1. Participants' subjective evaluation and eye movement data

Table 1. Design factors for each category of messages.

Message Categories	Design Factors
(1) Up-to-six-lines travel-time messages	• Five-line vs. six-line vs. three-line
(2) Transit travel-time messages	
– Bus	• Transit agency logo vs. transit agency logo with the bus symbol vs. pure text
– Light rail	• Transit agency logo vs. standard transit symbol
(3) GRIPs	
– Single-link GRIPs	• Orientation (top-bottom vs. bottom-top)With or without legend (i.e., road-work)
– Dual-link GRIPs	• Position of destinations (middle vs. left)

were collected. The dependent variables are (1) rating of easiness to understand; (2) understanding accuracy of general information; (3) understanding accuracy of detailed information; and (4) participants' glances behavior, including both gaze position and fixation time.

2.7. Pre-processing and statistical analysis of the data

Understanding accuracy of each message's general information was calculated based on two experimenters' debriefing and coding of whether each participant correctly answered the pre-defined comprehension questions (e.g., the message is about travel-time, the destination is downtown San Francisco, the transportation mode is taking BART).

Raw data from the Tobii eye tracker included the timestamp, x, and y position of each gaze point. The accuracy was 0.5–1 degree, translating to 50–100 pixels based on the resolution and dimension of the simulator screens and the average distance from participants' eyes to the screens. The number of gaze points for two participants was significantly smaller than others, which might be caused by the eye tracker's accuracy deterioration after calibration. These two participants' visual attention data were excluded from further analysis. The level of frequency of gaze points and the color codes are shown in Figure 8. These colors indicate attention from highest to lowest frequency with the same difference between each pair of adjacent colors. Gaze points were analyzed to show the distribution of visual attention onto different elements on each display. Fixation time, an indication of drivers' processing time, was also analyzed to compare the viewing time between different designs.

All dependent variables were analyzed to measure the understandability of the signs. Mixed ANOVA with the design factors as within-subject factor, gender, and language as between-subject factors was used to analyze the difference among different designs.

Figure 8. Color scale of the heatmap for visual attention data.

3. Results

3.1. Distribution of driving speed

Participants' driving speed before viewing the signs was recorded by the simulation program. To validate participants' immersion in driving during the experiments, the mean speed of each trial was calculated and shown in Figure 9. For the up to six lines travel-time messages and transit travel-time messages, in 77.24% of the trials, participants maintained the speed within the recommended speed range (55– 75 mph). For GRIPs, in 92.28% of the trials, participants maintained the speed within the recommended range (25– 45 mph). These results show that the participants well maintained the speed of the simulator as instructed. It indicates that participants were well-engaged in the simulated driving tasks.

3.2. Up-to-six-lines travel-time messages

3.2.1. Easiness to understand

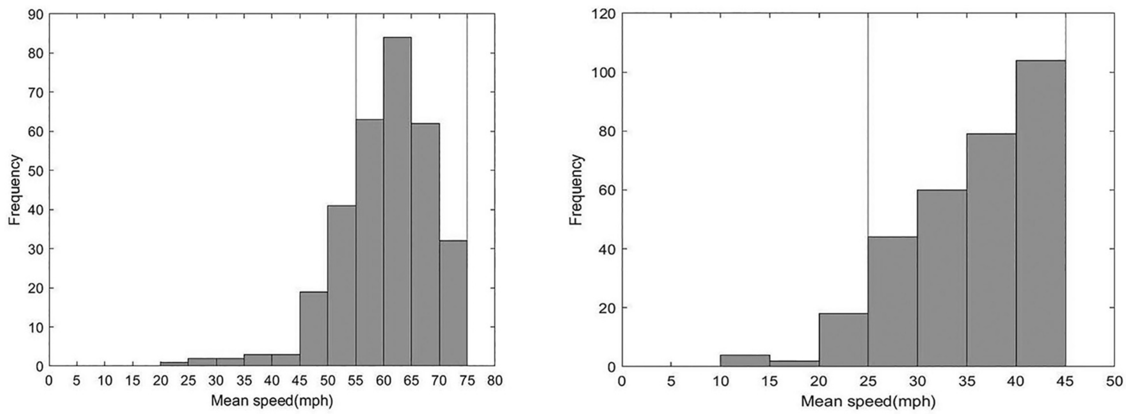
Mixed ANOVA was conducted to compare the mean of easiness among the three messages, in which gender and language were analyzed as between-subject factors. Table 2 shows the results of easiness, including the mean and standard deviation (SD) for each message. The mean of perceived easiness was 4.63 out of 5 (SD = 0.71) for the three-line message, 4.04 (SD = 1.08) for the five-line message, and 4.22 (SD = 1.09) for the six-line message. Results further indicated that the design factor had a significant impact on the perceived easiness of the three messages, $F(2, 40) = 4.30$, $p = .02$. Post-hoc analysis was conducted. Results indicated that the five-line message's perceived easiness was significantly lower than the three-line message, with a mean decrease of 0.59, $p = .012$. There was no significant difference between the three-line message and the six-line message. Test of between-subject effects showed that gender or language had no significant effect on the three messages' perceived easiness.

3.2.2. Understanding accuracy of general information

Mixed ANOVA was also used to compare the mean of understanding accuracy of general information. Similarly, gender and language were included as between-subject factors. Table 2 also shows the mean (SD) of understanding the accuracy of general information for each message. The average understanding accuracy of general information was 100% for both the three-line message and the five-line message. It was 90.28% (SD = 23.00%) for the six-line message. Results indicated that the design factor had no significant impact on understanding accuracy. However, there was a significant interaction between the design factor and language ($p = .038$). Results showed that, for native Spanish participants, the understanding accuracy of the six-line message (mean = 80.70%, SD = 0.30) was significantly lower than other conditions.

3.2.3. Understanding accuracy of detailed information

Table 2 also shows the understanding accuracy of detailed information, with the number (percentage) of the participants who correctly reported the information about the destination. For the five-line message, only 12 out of 24 participants (50%)

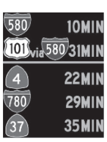


(a) Distribution of Speed for Travel-time Messages and Transit Travel-time Messages

(b) Distribution of Speed for GRIPs

Figure 9. Speed distribution before viewing the signs.

Table 2. Understandability and visual attention for travel-time and up-to-six lines messages.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Understanding accuracy of detailed information	Fixation time (second)	Heatmap of gaze points
1		4.63 * (0.71)	100.00% (0)	15 (62.50%)	11.24 (2.64)	
2		4.04 * (1.08)	100.00% (0)	12 (50.00%)	10.90 (3.24)	
3		4.22 (1.09)	90.28% (23.00%)	14 (58.33%)	10.00 (3.20)	

* Indicate significant difference, $p < .05$.

correctly found the information about the destination, which was the lowest among the three messages.

3.2.4. Visual attention

As shown in Table 2, participants focused their attention on each line of information for the three-line message. Besides, the word "via" on the third line drew more attention than other display elements. For the five-line message, participants didn't pay even attention to each line. Participants spent much time looking at the second line, which was related to their destination. They also spent much time on the information below the horizontal line, which was irrelevant to this task's destination. For the six-line message, the heatmap indicated that "San Jose" (destination of this trip) took a reasonable amount of attention. Similarly, the words "via" took participants' special attention compared to other display elements.

The distribution of fixation duration time for the travel-time messages is shown in Figure 10. It shows that the fixation time for the five-line message is more variable than the three-line message and the six-line message. It indicates that some participants found the relevant information to the destination right away, while some participants looked at other irrelevant information. Mixed ANOVA was used to explore the difference in fixation time. However, no significant difference was found.

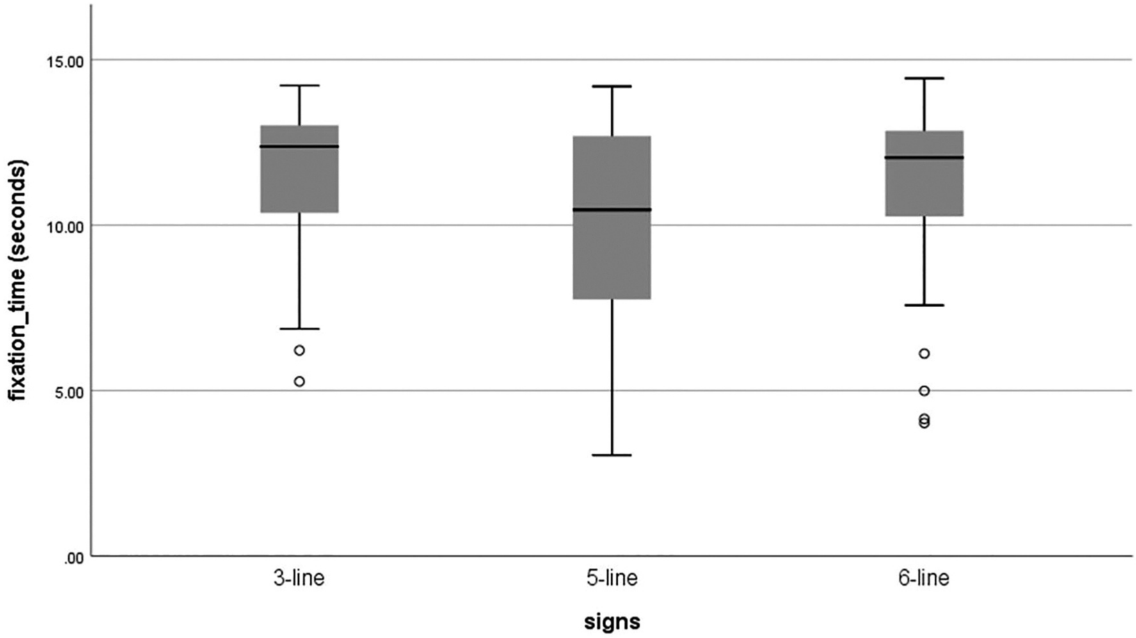


Figure 10. Fixation duration for the travel-time messages.

3.3. Transit travel-time messages

3.3.1. AC Transit

3.3.1.1. Easiness to understand. Table 3 shows the mean (SD) of the easiness ratings for the AC transit travel-time message for each design. The mean of perceived easiness was 4.17 (SD = 0.87) for the design with "AC Transit" logo, 3.98 (SD = 1.24) for the design with both "AC Transit" and bus symbols, and 4.06 (SD = 1.24) for the design with pure-text. Results indicated that the design factor had no significant impact on the perceived easiness of the three messages. Gender or language had no significant effect on the easiness to understand either.

3.3.1.2. Understanding accuracy of general information. Table 3 also shows the mean (SD) of understanding accuracy of general information for each design, which was 90.63% (SD = 14.39%), 83.33% (SD = 19.03%), and 89.58% (SD = 16.34%) respectively. However, no factor had a significant impact on understanding accuracy of general information.

Table 3. Understandability and visual attention for AC Transit travel-time message.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Understanding accuracy of detailed information	Perceived symbol helpfulness	Fixation time (second)	Heatmap of gaze points
1		4.17 (0.87)	90.63% (14.39%)	20 (83.33%)	4.08 * (1.24)	9.63 (2.93)	
2		3.98 (1.24)	83.33% (19.03%)	19 (79.17%)	2.65 * (2.31)	9.43 (3.54)	
3		4.06 (1.24)	89.58% (16.34%)	20 (83.33%)	NA	9.73 (2.47)	

* Indicate significant difference, $p < .05$.

3.3.1.3. Understanding accuracy of detailed information. Table 3 shows the understanding accuracy of detailed information. The understanding accuracy about the destinations for the three designs was similar to one another, which was 83.33% for the design with "AC Transit" logo, 79.17% for the design with both "AC Transit" logo and bus symbol, and 83.33% for the design with pure-text.

3.3.1.4. Helpfulness of symbols. Mixed ANOVA was also used to compare the mean of the perceived symbol helpfulness for the first and the second design of the AC transit travel-time messages. Table 3 shows the mean (SD) of the perceived symbol helpfulness for the two designs. Results indicated that the design factor significantly impacted the perceived symbol helpfulness, F (1, 19) = 6.81, p = .017. The symbol helpfulness of the design with the "AC Transit" logo (mean = 4.08, SD = 1.24) was significantly higher than the design with both the "AC Transit" logo and the bus symbol (mean = 2.65, SD = 2.31).

3.3.1.5. Visual attention. As shown in Table 3, for all three displays, participants paid attention to the first phase of the message (the starting point) and the second phase of the message (the destinations). For the 2nd display, participants paid little attention to the bus symbol. It indicated that most participants didn't see the bus symbol.

Fixation duration time for each AC Transit travel-time message is shown in Figure 11, which was averagely 9.63 seconds (SD = 2.93), 9.43 seconds (SD = 3.54), and 9.73 seconds (SD = 2.47) respectively. However, mixed ANOVA results indicated that the fixation time was not significantly different among different designs.

3.3.2. BART

3.3.2.1. Easiness to understand. Mixed ANOVA tests with one within-subject factor (i.e., design factor) and two between-subject factors (i.e., gender and language) were used to compare the alternative designs for the BART travel-time messages and the BART service special messages.

Table 4 shows the mean (SD) of easiness for each BART travel-time message design. It was 4.59 (SD = 0.65) for the symbol design, and 4.52 (SD = 0.66) for the pure-text design. Results indicated that the design factor had no significant impact on the perceived easiness of the two designs. However, there was a significant interaction effect between the design and gender, F (1, 20) =6.50, p = .019. Results showed that female participants rated the easiness of the pure-text design (mean = 4.00, SD = 1.47) significantly lower than their rating of the symbol design and male participants' rating for both designs.

Table 5 shows the mean (SD) of easiness for each BART service special message design. It was 4.46 (SD = 1.14) for the light-rail symbol design, and 4.92 (SD = 0.28) for the BART-logo design. Results indicated that the design factor had no significant impact on the perceived easiness. Results also showed that gender or language had no significant effect on the perceived easiness either.

3.3.2.2. Understanding accuracy of general information. Table 4 shows the mean (SD) understanding accuracy of general information for the BART travel-time messages, which was 95.83% (SD = 9.54%) for both designs. Table 5 shows the mean (SD) understanding accuracy of general information for the BART service special messages, which was 97.22% (SD = 9.43%) and 98.61% (SD = 6.78%) respectively. However, the design factor had no significant impact on understanding accuracy of general information. Test of between-subject effects showed that gender or language had no significant effect either.

3.3.2.3. Understanding accuracy of detailed information. Table 4 also shows the understanding accuracy about the destination for the BART travel-time message. The understanding accuracy about the destination for the first design (with BART logo) (87.50%) was higher than the second design (with text only) (58.33%).

3.3.2.4. Helpfulness of symbols. Table 5 shows the mean (SD) of perceived symbol helpfulness for the BART service special message. Results indicated that the symbol helpfulness for the BART logo (mean = 4.96, SD = 0.20) was significantly higher than the light-rail symbol (mean = 4.35, SD = 0.81), F (1, 20) = 12.94, p = .002.

3.3.2.5. Visual attention. Heatmaps for the BART travel-time messages are also shown in Table 4. Participants paid attention to both the first phase of the message (the starting point) and the second phase of the message (the destinations) for the two displays. In the two heatmaps, special attention was on the word "STATION," which indicated that "STATION" was the critical word in the message for participants to understand the first phase (the starting point).

Heatmaps for the BART special messages are shown in Table 5. For the two displays, participants paid attention to the BART symbol and the word "NO," which was the message's essential information. The light-rail symbol on the first display captured more attention than other elements on the sign, implying that the light-rail symbol took more time for the drivers to comprehend.

The fixation duration time for each BART message is shown in Figure 12. Mixed ANOVA was conducted to find the influence on fixation time. However, no significant difference was found, either for the BART travel-time message or for the BART special message.

3.4. GRIPs

3.4.1. Single-link GRIPs

3.4.1.1. Easiness to understand. Mixed ANOVA was conducted to determine the effects of the two design factors (i.e., orientation, legend) on easiness ratings of the four GRIPs. Gender and language were included as between-subject factors. Table 6 shows the results of the mean (SD) of easiness for each message. It was 4.38 (SD = 1.05) for the bottom-up without legend design, 4.38 (SD = 1.10) for the top-down without legend design, 4.71 (SD = 0.62) for the bottom-up with legend design, and 4.15 (SD = 1.24) for the top-down with legend design. Results indicated that the orientation design had no significant impact on the perceived easiness of the four displays. Results also

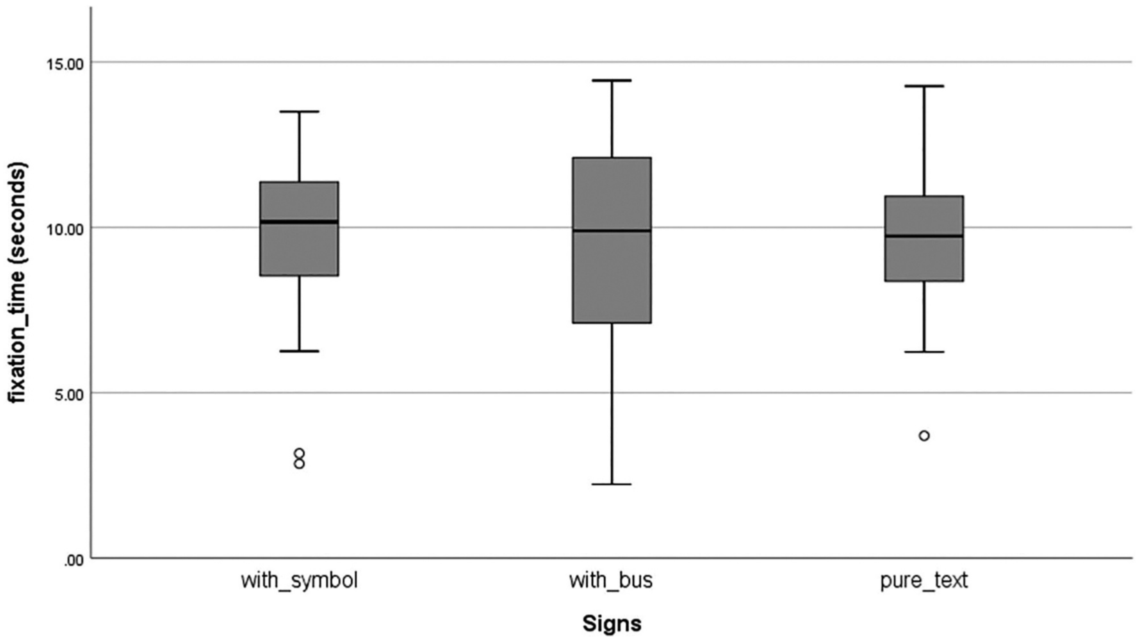


Figure 11. Fixation duration for the ac transit travel-time messages.

Table 4. Understandability and visual attention of BART travel-time message.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Understanding accuracy about destinations	Fixation time (second)	Heatmap of gaze points
1		4.59 * (0.65)	95.83% (9.54%)	21 (87.50%)	10.93 (2.49)	
2		4.52 * (0.66)	95.83% (9.54%)	14 (58.33%)	10.77 (2.89)	

* Indicate significant difference, $p < .05$.

Table 5. Understandability and visual attention of BART special message.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Perceived symbol helpfulness	Fixation time (second)	Heatmap of gaze points
1		4.46 (1.14)	97.22% (9.43%)	4.35 * (0.81)	7.25 (3.39)	
2		4.92 (0.28)	98.61% (6.78%)	4.96 * (0.20)	8.52 (3.29)	

* Indicate significant difference, $p < .05$.

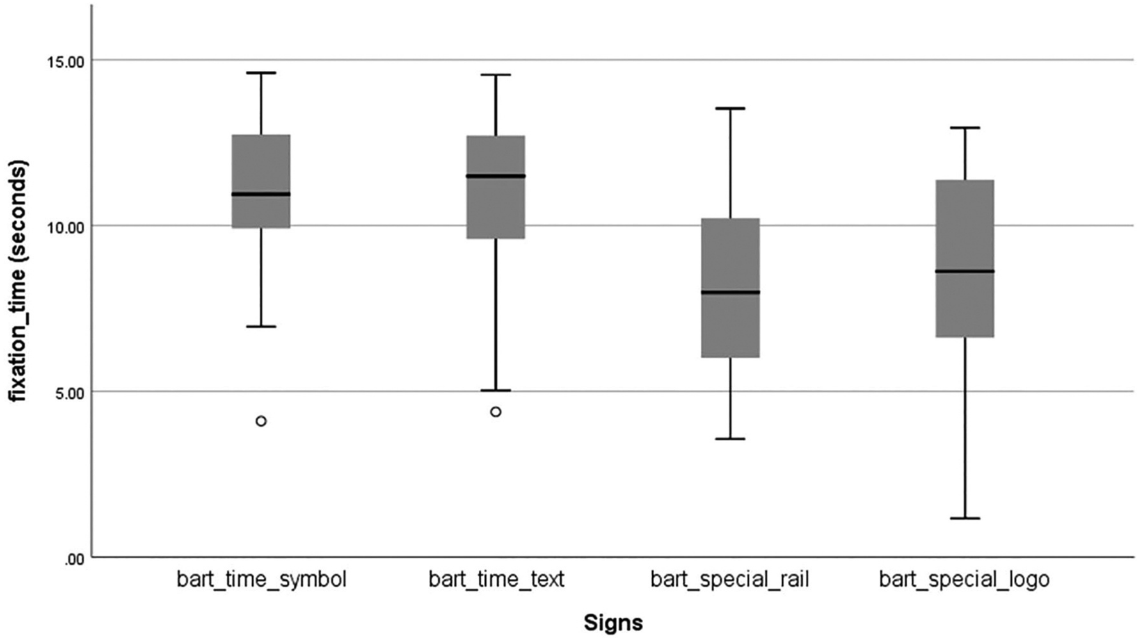


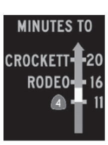
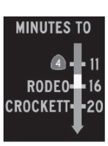
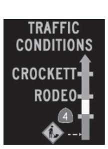
Figure 12. Fixation duration for BART messages.

indicated that having legend had no significant effect on easiness either. Test of between-subject effects showed that gender or language had no significant effect on the easiness to understand.

3.4.1.2. Understanding accuracy of general information.

Mixed ANOVA was also used to determine the effects of the two design factors on understanding accuracy of general information. Table 6 shows the results of the mean (SD) for

Table 6. Understandability and visual attention of single-link GRIPs.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Understanding accuracy of detailed information	Fixation time (second)	Heatmap of gaze points
1		4.38 (1.05)	100.00% * (0)	Orientation: 23 (95.83%)	12.71 * (3.52)	
2		4.38 (1.10)	100.00% * (0)	Orientation: 23 (95.83%)	12.99 * (3.34)	
3		4.71 (0.62)	98.96% * (5.10%)	Legend: 24 (100.00%)	13.56 * (4.02)	
4		4.15 (1.24)	96.88% * (8.45%)	Legend: 24 (100.00%)	14.33 * (3.22)	

* Indicate significant difference, $p < .05$.

each message. Similarly, the design of orientation had no significant impact on understanding accuracy of general information. For the displays without legend, the mean understanding accuracy (mean = 100.00%, SD = 0) was higher than the displays with legend (mean = 98.96%, SD = 5.10%; mean = 96.88%, SD = 8.45%). However, results indicated that the legend had no significant effect either. Gender or language had no significant effect on understanding accuracy of general information.

3.4.1.3. Understanding accuracy of detailed information. Table 6 also shows the understanding accuracy of detailed information (e.g., understanding of orientation, the legend). The understanding accuracy of the orientation for the first and second designs was the same (95.83%). The understanding accuracy of the legend for the third and fourth designs was also the same (100.00%).

3.4.1.4. Visual attention. According to the heatmap of gaze points, participants paid the most attention to the link and the destinations. For example, for the first display shown in Table 6, participants looked at all three locations (Crockett, Rodeo, and Highway 4) and spent more attention on the destination of the task (Crockett). Similarly, two sets of mixed ANOVAs were conducted to explore the difference in fixation time. It was found that there was a significant difference between displays with legend and displays without legend, F (1, 21) = 6.88, p = .016. Participants looked at the displays with legend longer than the displays without legend.

3.4.1.5. Comparison between three-line travel-time message and single-link GRIPs. Mixed ANOVAs were conducted to compare the mean of easiness and the mean of understanding accuracy of the three-line message and the four single-link-GRIP messages, in which gender and language were analyzed as between-subject factors. However, there is no statistically significant difference either for the easiness or the understanding accuracy among the four messages. Neither gender nor language have a significant effect on easiness and understand accuracy. In the experiment set-up, the total viewing time for GRIP was 20 seconds and 15 seconds for the three-line travel time messages. Hence, the fixation time for the three-line message and the four single-link-GRIP messages was not compared.

3.4.2. Dual-link GRIPs

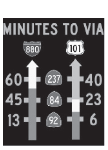
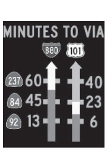
3.4.2.1. Easiness to understand. Mixed ANOVA was conducted to determine the effect of the position of destinations for the two Dual-link GRIPs. Gender and language were included as two between-subject variables. As shown in Table 7, the position of destinations had no significant effect on perceived easiness. The mean of perceived easiness was 4.21 (SD = 1.06) for the design with destinations in the middle of the two links, and was 4.27 (SD = 0.64) for the design with destinations on the left side of the two links. Gender or language had no significant effect either.

3.4.2.2. Understanding accuracy of general information. Mixed ANOVA was also used to analyze the understanding accuracy of general information. The mean (SD) understanding accuracy is also shown in Table 7, which was 98.61% (SD = 6.8%) for both designs. Results indicated that the position of destinations had no significant effect on the understanding accuracy of general information.

3.4.2.3. Understanding accuracy of detailed information. Participants were asked about their understanding of the destinations (i.e., highway 237, 84, 92) and ratings of the easiness to tell those destinations. For the first display, 22 (91.67%) of participants understood the destinations correctly. It was higher than the second display, in which 17 (70.83%) participants had understood it correctly. Mixed ANOVA was conducted to test the difference of easiness rating for the two positions of destinations. However, no significant difference was found.

3.4.2.4. Comparison between single-link and dual-link GRIPs. Mixed ANOVAs were conducted to compare the mean of easiness, the mean of understanding accuracy, and the mean of fixation time for the single-link and dual-link GRIPs, in which gender and language were analyzed as between-subject factors. For the easiness and the understanding accuracy, there is no statistically significant difference among the six messages.

Table 7. Understandability of dual-link GRIPs.

	Graphical Images	Perceived easiness	Understanding accuracy of general information	Understanding accuracy of detailed information	Position of destinations	Fixation time (second)	Heatmap of gaze points
1		4.21 (1.06)	98.61% (6.8%)	22 (91.67%)	3.91 (1.44)	15.41 (3.68)	
2		4.27 (0.64)	98.61% (6.8%)	17 (70.83%)	4.39 (1.08)	15.53 (2.84)	

The distribution of fixation time for both single-link GRIPs and dual-link GRIPs is shown in Figure 13. The mixed ANOVA results for fixation time show a significant difference among the six GRIP messages ($F = 5.46$, $p < .001$). Pairwise comparison results indicate that the fixation time between the single-link GRIP with bottom-top orientation without a legend and the dual-link GRIP with destinations in the middle has a significant difference of 2.7 seconds ($p = .015$); the single-link GRIP with top-bottom orientation without a legend and the dual-link GRIP with destinations in the middle has a significant difference of 2.4 seconds ($p = .049$), and a significant difference of 2.5 seconds ($p = .034$) with the dual-link GRIP with the destination on the left. Neither gender nor language have a significant effect on easiness and understand accuracy.

4. Discussion

4.1. Message length

To determine the effect of message length on drivers' comprehension of the IDB signs, we used the three-line message as the baseline and compared participants' subjective evaluation and eye movement data for the five-line message and six-line message. Results indicated that participants rated the five-line message as more difficult to understand than the three-line message. On the other hand, there was no significant difference in easiness between the three-line and six-line messages. Besides, the five-line message's visual attention data indicated that participants spent a meaningful amount of attention on irrelevant information. One possible reason could be that the layout of the five-line message (with two lines on the top and three lines at the bottom divided by the virtual line) was not intuitive, which took drivers extra attention to figure that out. Another finding was that there was a significant interaction between the design factor and the driver's native language, which showed that native Spanish participants had significantly lower understanding accuracy for the six-line message than the native English participants. It implies that when a sign's complexity increases, non-native English drivers are more likely to have lower comprehension accuracy of the messages.

In addition, heatmap of eye movement data shows that, for the five-line message, participants paid attention to each line on the sign, while only one line of the message is related to the destination of the current trial. For the six-line message, participants figured out the message's layout and omitted the irrelevant lines of information on the sign. As a result, most of them only paid attention to four lines out of the six-line message, which could explain why the overall easiness rating for the six-line message is not significantly different from the three-line message. While for the five-line message, some participants took much longer visual attention to the whole sign. As a result, the distribution of the fixation time for the five-line message is much spread out. The results could imply that the layout of dividing the five-line messages into two virtual phases doesn't make sense and induced confusion to the drivers instead.

As mentioned in the literature review, existing research (Dudek & Ullman, 2006; Dudek et al., 2007) and the engineering practice of CMS message design follow the guideline that three units of information should be displayed on a CMS. Based on the current study results, this guideline is also suggested to be applied for advanced CMS with higher resolution LED displays, although they are capable of displaying more units of information. However, when the message could be divided into two virtual phases parallel to each other but not inter-related, as in the six-line message, drivers could easily find relevant lines to look into and omit the irrelevant lines.

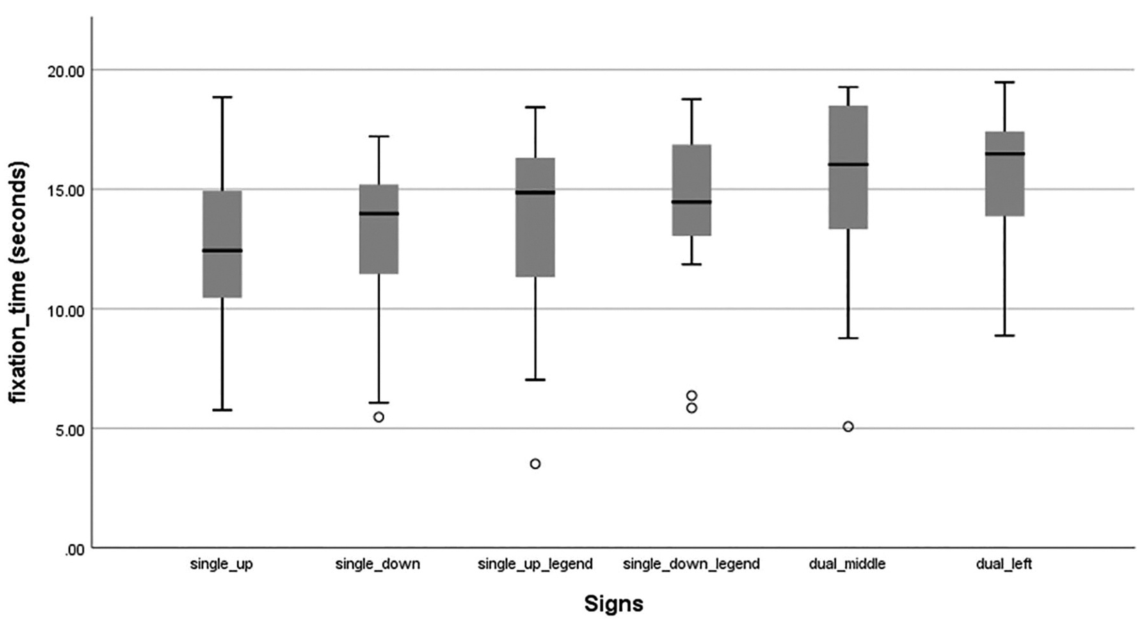


Figure 13. Fixation duration for the GRIPs.

4.2. Use of graphics

In this study, firstly, we found that the IDB signs with symbols or transit logos were more preferred and rated as easier to understand than those pure-text signs, which is well-aligned with most existing studies (e.g., Ells & Dewar, 1979). For the AC transit travel-time messages, results indicated that having only the AC transit logo on top of the sign was more helpful than having both the AC transit logo and a bus symbol. All participants were local commuters. Based on their comments during the experiment, most of them were familiar with the AC transit before attending this study. When a bus symbol was used in addition to the AC transit logo, it was perceived as redundant and not helpful. Hence, drivers' familiarity with the signs played an important role in the perception of the bus symbol. As found from the study by Ben-Bassat and Shinar (2006), the comprehension of a traffic sign is highly related to drivers' familiarity with relevant information from their own driving experience. Based on the findings of this study, it is recommended that the design of the symbol should also consider drivers' familiarity with the corresponding objects (i.e., AC transit, BART).

The results of the BART travel-time message are similar. Particularly, female participants rated the signs with symbols easier to understand than the signs with only text. Almost all participants preferred to have the BART logo for the BART special message rather than the generic light-rail symbol. One reason is that the light-rail symbol could mean other train services (e.g., Cal-train, Amtrak). That is to say, the light-rail symbol is not specific enough for BART. Another reason is that all participants of this study were local commuters who knew the BART logo very well. For the light-rail symbol recommended by CA MUTCD (California Department of Transportation, 2015), drivers see it much less frequently. As the objective of the IDBs is to provide traveler information for local commuters, the well-known BART logo is preferred over the generic light-rail symbol. Yet, it proves that the driver's familiarity with the elements of the message plays an important role in sign comprehension.

4.3. GRIPs

For the single-link GRIPs, the orientation does not have an impact on participants' comprehension. However, having the legend takes significantly longer visual attention of the participants. On the other hand, the legend does not have a significantly positive effect on the understandability of the single-link GRIPs. Therefore, adding a legend for single-link GRIP is not recommended. A comparison between single-link GRIPs and the baseline three-line message shows that the understanding accuracy of the single-link GRIP (with 3 destinations) is not significantly different from the three-line message. This result implies that single-link GRIPs with 3 destinations have similar understandability as the three-line message. As single-link GRIPs provide more information (congestion levels) than the three-line message, they are recommended for further research and implementation. Results showed that all participants understood the color codes (i.e., green, yellow, and red) to indicate congestion levels.

Another finding for the single-link GRIP is the learning effect, which is noticeable between the two stages of this project. As shown in P. Wang et al. (2019), the understanding accuracy in the initial stage ranges from 83% to 91%. In this simulator testing, many participants mentioned that it was much easier to understand the single-link GRIPs. Meanwhile, the understanding accuracy for single-link GRIPs ranges from 96.88% to 100%. Alkim (2000) pointed out that "comprehension of the information presented is initially quicker for regular CMS. But drivers get used to GRIPs rapidly, and the difference is made up quickly." In addition, the GRIPs are proposed to be only displayed during congested periods when drivers have repeated and extended opportunities to view the displays. Therefore, the GRIPs have the potential to provide more comprehensive traveler information on CMS with satisfactory understandability.

For the dual-link GRIPs, the visual attention data shows that, on average, it took around 15 seconds for participants to comprehend the information on the signs. They also took significantly longer visual attention than the single-link GRIPs. While answering the post-trial questions, many participants commented that they were overwhelmed by the amount of information shown on the dual-link GRIPs. Therefore, they are recommended to be further revised and evaluated before being implemented.

4.4. Drivers' difference

In this study, we investigated two driver-related factors: gender and native language. It was found that the interaction with the design factor and the language played a role in comprehending certain signs. We also found that female participants rated the design with symbols more helpful than the design with pure text for the BART travel-time message. But for understanding accuracy, there was no significant difference between female and male participants. These findings provide additional evidence to the existing literature (Al-Madani & Al-Janahi, 2002) that driver-related factors such as gender and language played an important role in comprehending traffic signs.

4.5. Limitations of this study and future research directions

There were several limitations of this study. Firstly, this study only focused on the comprehension of the signs with 24 participants, which may raise generalizability concerns. However, the driving simulator study reported here was one part of a bigger project. In this project, multiple phases of human-factors studies have been proposed to do a comprehensive evaluation of different IDB signs. The first phase of this project was a lab-testing, with 50 local commuters participated. It mainly focused on the assessment of static comprehension (P. Wang et al., 2019). The simulator testing reported in this paper was the second phase, which focused on evaluating dynamic comprehension. In the third phase of the project, we proposed to evaluate the legibility distance of the signs in different traffic and weather conditions. With the three phases of studies, we aimed to provide recommendations for a safe real-world demonstration with precautions that certain categories of signs are easy to understand and have potential benefits while some others may be difficult to understand and more problematic.

5. Conclusions

As a summary, this current study sheds light on various aspects of the message design for CMSs using advanced display technologies. Based on the results of the current study, we recommend that advanced CMS with higher resolution LED displays should also follow the guideline of displaying three lines of information on one display. When the objective of the CMS is to provide traveler information for local commuters, drivers' familiarity with the objects (e.g., bus, BART) should be considered for selecting or designing the symbols. The well-known symbols (e.g., BART logo) should be used over other standards symbols (e.g., light-rail). This finding should apply to other locally well-known brands of light-rail transit services as well. Single-link GRIPs with 3 destinations have similar understandability as the three-line message. As single-link GRIPs also provide more information (congestion level) than the three-line message, they are recommended for further research and implementation. However, adding more information onto the sign, such as the road-work legend or incident legend, will increase the complexity of the sign, which is not recommended. For the same reason, dual-link GRIPs with three interchanges or destinations are not recommended either. In the future, on-road experiments to collect field data will likely provide additional insights regarding the overall performance of the IDBs, allowing researchers to evaluate the legibility distance of these non-approved IDB displays.

Acknowledgments

This work was supported by the California Department of Transportation (Caltrans). Their support throughout this project is highly appreciated. The views and opinions expressed in this article are those of the authors and do not necessarily reflect Caltrans and associated agencies' official policy or position.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alkim, T., & Schenk, B. (2001, September 30 - October 4). Graphical route information. *8th World Congress on Intelligent Transport Systems, Sydney, Australia.*

Alkim, T. P. (2000). Graphical route information on variable message signs. *Tenth International Conference on Road Transport Information and Control*, 32–36. https://doi.org/10/b6kqxp

Alluri, P., Gan, A., Saha, D., & Fernandez, L. O. (2017). Evaluation of signage alternatives for express lane facilities: Focus group study. *Transportation Research Record: Journal of the Transportation Research Board*, *2616*(1), 114–121. https://doi.org/10.3141/2616-12

Al-Madani, H., & Al-Janahi, A.-R. (2002). Assessment of drivers' comprehension of traffic signs based on their traffic, personal and social characteristics. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *5*(1), 63–76. https://doi.org/10/btt8zm

Ben-Bassat, T., & Shinar, D. (2006). Ergonomic guidelines for traffic sign design increase sign comprehension. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(1), 182–195. https://doi.org/10.1518/001872006776412298

Brockmann, R. J. (1991). The unbearable distraction of color. *IEEE Transactions on Professional Communication*, *34*(3), 153–159. https://doi.org/10.1109/47.84109

California Department of Transportation. (2014). California manual on uniform traffic control devices.

California Department of Transportation. (2015). California MUTCD 2014 edition. Chapter 2L. Changeable message signs.

Campbell, J. L., Carney, C., & Kantowitz, B. H. (1998). *Human factors design guidelines for advanced traveler information systems (ATIS) and commercial vehicle operations (CVO) (FHWA-RD-98-057)*. FHWA. https://www.fhwa.dot.gov/publications/research/safety/98057/toc.cfm

Charlton, S. G. (2006). Conspicuity, memorability, comprehension, and priming in road hazard warning signs. *Accident Analysis and Prevention*, *38*(3), 496–506. https://doi.org/10.1016/j.aap.2005.11.007

Crundall, D., Crundall, E., Burnett, G., Shalloe, S., & Sharples, S. (2011). The impact of map orientation and generalisation on congestion decisions: A comparison of schematic-egocentric and topographic-allocentric maps. *Ergonomics*, *54*(8), 700–715. https://doi.org/10.1080/00140139.2011.592608

Dewar, R., Kline, D., & Swanson, H. (2001). Age differences in comprehension of traffic symbols. *Accident Reconstruction Journal*, *12*(1), 37–41. http://onlinepubs.trb.org/Onlinepubs/trr/1994/1456/1456-001.pdf

Dewar, R. (1990). Criteria for the design and evaluation of traffic sign symbols. *Applied Ergonomics*, *21*(1), 79–80. https://doi.org/10/dnc5c8

Diop, E. B., Zhao, S., & Duy, T. V. (2019). An extension of the technology acceptance model for understanding travelers' adoption of variable message signs. *PLoS ONE*, *14*(4), p. e0216007. https://doi.org/10/ghnr2s

Dudek, C. (2004). *Changeable message sign operation and messaging handbook*. Texas Transportation Institute. https://tmcpfs.ops.fhwa.dot.gov/cfprojects/uploaded_files/CMS%20Operation%20and%20Messaging%20Handbook-Final%20Draft.pdf

Dudek, C. L., & Ullman, G. L. (2006). *Dynamic message sign message design and display manual* (Report No. FHWA/TX-04/0-4023-P3). Texas Transportation Institute. https://static.tti.tamu.edu/tti.tamu.edu/documents/0-4023-P3.pdf

Dudek, C. L., Schrock, S. D., & Ullman, B. R. (2007). License plate and telephone numbers in changeable message sign amber alert messages. *Transportation Research Record: Journal of the Transportation Research Board*, *2012*(1), 64–71. https://doi.org/10/fd9rvv

Dutta, A., Fisher, D. L., & Noyce, D. A. (2004). Use of a driving simulator to evaluate and optimize factors affecting understandability of variable message signs. *Transportation Research Part F-Traffic Psychology and Behaviour*, *7*(4–5), 209–227. https://doi.org/10/dtrf69

Ells, J. G., & Dewar, R. E. (1979). Rapid comprehension of verbal and symbolic traffic sign messages. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *21*(2), 161–168. https://doi.org/10.1177/001872087902100203

FHWA. (2009). *Manual on uniform traffic control devices 2009 edition including revisions 1 and 2 in May 2012* (2009th ed.). US Department of Transportation.

Fisher, J. (1992). Testing the effect of road traffic signs′ informational value on driver behavior. *Human Factors*, *34*(2), 231–237. https://doi.org/10/ghnrhd

Gan, H., Sun, L., & Chen, J. (2008). Advanced traveler information system for metropolitan expressways in Shanghai. *8th International Conference of Chinese Logistics and Transportation Professionals - Logistics: The Emerging Frontiers of Transportation and Development in China*, 343(1), 1810–1818. https://doi.org/10.1061/40996(330)265

Gan, H., & Ye, X. (2013). Investigation of drivers' diversion responses to urban freeway variable message signs displaying freeway and local street travel times. *Transportation Planning and Technology*, *36*(8), 651–668. https://doi.org/10/ghnr3v

Gao, X. W., Podladchikova, L., Shaposhnikov, D., Hong, K., & Shevtsova, N. (2006). Recognition of traffic signs based on their colour and shape features extracted using human vision models. *Journal of Visual Communication and Image Representation*, *17*(4), 675-685. https://doi.org/10/btwn7t

ISO. (2011). *ISO 3864-1:2011—Graphical symbols—Safety colours and safety signs—Part 1: Design principles for safety signs and safety markings*. https://www.iso.org/standard/51021.html

Jamson, S. L., Tate, F. N., & Jamson, A. H. (2005). Evaluating the effects of bilingual traffic signs on driver performance and safety. *Ergonomics*, *48*(15), 1734–1748. https://doi.org/10.1080/00140130500142191

Lai, C.-J. (2010). Effects of color scheme and message lines of variable message signs on driver performance. *Accident Analysis & Prevention*, *42*(4), 1003–1008. https://doi.org/10.1016/j.aap.2009.12.002

Lai, C.-J. (2012). Drivers' comprehension of traffic information on graphical route information panels. *Accident Analysis & Prevention*, *45*, 565–571. https://doi.org/10.1016/j.aap.2011.09.008

Macdonald, W. A., & Hoffmann, E. R. (1991). Drivers' awareness of traffic sign information. *Ergonomics*, *34*(5), 585–612. https://doi.org/10/fpwt5t

Mackett-Stout, J., & Dewar, R. (1981). Evaluation of symbolic public information signs. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *23*(2), 139–151. https://doi.org/10.1177/001872088102300202

Martens, M. H. (2000). Assessing road sign perception: A methodological review. *Transportation Human Factors*, *2*(4), 347–357. https://doi.org/10.1207/STHF2-4_4

Campbell, J. L., & National Research Council (US), National Cooperative Highway Research Program, American Association of State Highway and Transportation Officials, & United States (Eds.). (2012). *Human factors guidelines for road systems*. 2nd ed. Transportation Research Board.

Ng, A. W. Y., & Chan, A. H. S. (2007). The guessability of traffic signs: Effects of prospective-user factors and sign design features. *Accident Analysis and Prevention*, *39*(6), 1245–57. https://doi.org/10/fkf88g

Ou, Y.-K., & Liu, Y.-C. (2012). Effects of sign design features and training on comprehension of traffic signs in Taiwanese and Vietnamese user groups. *International Journal of Industrial Ergonomics*, *42*(1), 1–7. https://doi.org/10.1016/j.ergon.2011.08.009

Richards, A., Mcdonald, M., Fisher, G., & Brackstone, M. (2004). Investigation of Driver Comprehension of Traffic Information on Graphical Congestion Display Panels using a Driving Simulator. *European Journal of Transport and Infrastructure Research*, *4*(4), 417–435.

Roca, J., Tejero, P., & Insa, B. (2018). Accident ahead? Difficulties of drivers with and without reading impairment recognising words and pictograms in variable message signs. *Applied Ergonomics*, *67*, 83–90. https://doi.org/10.1016/j.apergo.2017.09.013

Roca, J., Insa, B., & Tejero, P. (2018). Legibility of text and pictograms in variable message signs: Can single-word messages outperform pictograms? *Human Factors*, *60*(3), 384–396. https://doi.org/10/ghnr2x

Schonfeld, G., Reischl, A., & Tsavachidis, M. (2000). Dynamic driver information goes graphical – A new quality in urban traffic information. In *Proceedings of World Congress on Intelligent Transport Systems (7th:2000), Turin, Italy*.

Shinar, D., & Vogelzang, M. (2013). Comprehension of traffic signs with symbolic versus text displays. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *18*, 72–82. https://doi.org/10.1016/j.trf.2012.12.012

Takeda, H., & Kujirai, S. (1999, November 8-12). Upgrading of information sign system on metropolitan expressway. *The 6th World Congress on Intelligent Transport Systems, Toronto, Canada*.

Ullman, B. R., Trout, N. D., & Dudek, C. L., & Institute, T. T., of Transportation, T. D., & Administration, F. H. (2009). *Use of graphics and symbols on dynamic message signs: Technical report*. 7(2), 192. https://trid.trb.org/view/890543

Wang, J. H., Collyer, C. E., & Yang, C. M. (2006). *Enhancing motorist understanding of variable message signs* (FHWA-RIDOT-RTD-06-1). University of Rhode Island. http://www.dot.ri.gov/documents/about/research/RIDOT_RTD_06-1.pdf

Wang, P., Zhang, T., & Chan, C.-Y. (2019, January 13-17). Evaluation of messages on information display board (IDB) using a laboratory experiment. *Transportation Research Board 98th Annual Meeting, Washingto DC, USA*. https://www.researchgate.net/publication/347074432_Evaluation_of_Messages_on_Information_Display_Board_IDB_Using_Laboratory_Experiment

Xu, C., Wu, Y., Rong, J., & Peng, Z. (2020). A driving simulation study to investigate the information threshold of graphical variable message signs based on visual perception characteristics of drivers. *Transportation Research Part F-Traffic Psychology and Behaviour*, *74*, 198–211. https://doi.org/10/ghnrz7

Yan, X., & Wu, J. (2014). Effectiveness of variable message signs on driving behavior based on a driving simulation experiment. *Discrete Dynamics in Nature and Society*, *2014*, 1–9. https://doi.org/10/gbfvj5

Zhao, W., Quddus, M., Huang, H., Lee, J., & Ma, Z. (2019). Analyzing drivers' preferences and choices for the content and format of variable message signs (VMS). *Transportation Research Part C-Emerging Technologies*, *100*, 1–14. https://doi.org/10/ghnr2t

About the Authors

Pei Wang is a research scientist at the California PATH program since June 2017. Her research focus is human factors in the transportation domain. Pei had been working in General Motors Research & Development and has a Ph.D. degree in Human Factors and Ergonomics from the Tsinghua University in China.

Tingting Zhang is an associate specialist at the California Partners for Advanced Transportation Technology program since November 2017. Before joining PATH, Zhang had been working in General Motors Research & Development as a human machine interaction researcher. She received her B.S degree from Guangdong Sun Yat-Sen University, China in 2009.

Xiao Zhou is an Associate Professor in Mechanical and Electrical School, Wuhan University of Technology, Wuhan, China. He received his Ph.D degree and did postdoc in Huazhong University of Science and Technology (HUST), Wuhan, China. His research interests are machine vision and intelligent system.

Sanaz Motamedi is a faculty member at Industrial and System Engineering, the University of Florida after her postdoctoral fellowship at the California Partners for Advanced Transportation Technology program, University of California, Berkeley. Her research interests are human factors, user experience, engineering education, and transportation safety.

Ching-Yao Chan is a Research Engineer and the Program Leader at California PATH. He is also serving as Associate Director of Berkeley Deep Drive. He has 30 years of research experience in a broad range of automotive and transportation systems. He obtained his doctoral degree from UC Berkeley in 1988.