Enhancing Driving Safety through Proper Message Design on Variable Message Signs

Jyh-Hone Wang, Ph.D.
Charles E. Collyer, Ph. D.
Chun-Ming Yang
University of Rhode Island

September 2005

URITC PROJECT NO. 000052

PREPARED FOR

UNIVERSITY OF RHODE ISLAND TRANSPORTATION CENTER

1. Report No	2. Government Accession No.	3. Recipient's Catalog No.
URITC FY 03-03	N/A	N/A

4. Title and Subtitle	5. Report Date
Enhancing Driving Safety through Proper Message Design on Variable Message Signs	September 2005
	6. Performing Organization Code N/A
7. Authors(s) Jyh-Hone Wang, Ph.D., Charles E. Collyer, Ph.D., and Chun-Ming Yang	8. Performing Organization Report No. N/A
9. Performing Organization Name and Address University of Rhode Island, Dept. of Industrial and Manufacturing Engineering, Kingston, RI 02881 (401) 874-5195 jhwang@egr.uri.edu	10. Work Unit No. (TRAIS) N/A
	11. Contract or Grant No. URI 500-2305-0000-0000052
	13. Type of Report and Period Covered Final Report, July 2003 – July 2005
12. Sponsoring Agency Name and Address University of Rhode Island Transportation Center Carlotti Administration Building, 75 Lower College Road Kingston, RI 02881	14. Sponsoring Agency Code A study conducted in cooperation with U.S. DOT

15. Supplementary Notes

N/A

16. Abstract

This report presents a study that assessed drivers' responses to and comprehension of variable message sign (VMS) messages displayed in different ways with the intent to help enhance message display on VMSs. Firstly, a review of literatures and current practices regarding the design and display of VMS messages is presented. Secondly, the study incorporates three approaches in the assessment. Questionnaire surveys were designed to investigate the preferences of highway drivers in regards to six message display settings, they were: number of message frames, flashing effect, color, color combinations, wording, and use of abbreviations. Lab experiments were developed to assess drivers' responses to a variety of VMS messages in a simulated driving environment. Two groups of factors, within-subject and between-subject factors, were considered in the design of experiment. Within-subject factors included message flashing and color combination. Between-subject factors were age and gender. To help validate results found from lab experiments, field studies were set up to study drivers' response to VMS in real driving environment. Thirty-six subjects, from three age populations (20-40, 40-60, above 60 yrs old) with balanced genders, were recruited to participate in both questionnaire surveys and lab experiments while eighteen of them participated in field studies on a voluntarily basis. The study findings suggest a specific set of VMS features that might help traffic engineers and highway management design VMS signs that could be noticed, understood and responded to in a more timely fashion. Safer and more proactive driving experiences could be achieved by adopting these suggested VMS features.

17. Key Words	18. Distribution Statement
Variable message signs, human factors, design of experiment.	No restrictions. This document is available to the Public through the URI Transportation Center, Carlotti Administration Building, 75 Lower College Rd., Kingston, RI 02881

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	64	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized (art. 5/94)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Variable message sign (VMS) technology has been increasingly used by highway authorities as a key component in intelligent transportation systems (ITS). It has become an effective tool not only to alleviate acute traffic problems, but also to help enhance driving safety. Also known as dynamic message sign (DMS), or changeable message sign (CMS), VMS is a programmable electronic bulletin board capable of displaying messages composed of text, pictogram, or both. Permanently mounted overhead VMSs are usually employed to convey pertinent, real-time traffic information that can be updated remotely. Portable VMSs are used mostly in construction, work zones, or special events to provide timely guidance to drivers.

With recent advancements in VMS technology, the permanently mounted overhead system is now equipped with a full-matrix, tri-color/full-color display that offers highway authorities great flexibility to post messages utilizing combinations of various font sizes and colors. In addition, dynamic capabilities such as alternating, flashing, scrolling, and even animation, have become standard features on a VMS system. These new features enable a VMS message to be displayed in a single frame or multiple frames, with or without flashing effect, etc. If desired, the message could also be shown in some combination of these dynamic effects.

To provide the public with a more efficient and safer daily travel environment, the highway authority in Rhode Island (RI) has deployed a number of state-of-the-art VMSs to communicate instantaneous traffic information and travel advice to drivers since late 2002. At present, there are a total of thirteen permanently mounted overhead VMSs in operation on major highways, such as Interstate Route 95, 195, 295 and RI Route 4. The systems installed in RI are the Vanguard® VF-2000 series manufactured by Daktronics (*1*). Each individual system has a full-matrix, tri-color display and is permanently mounted on an overhead sign bridge. The approximate sign dimensions are 8534 mm (28 feet) wide by 2388 mm (7 feet 10 inches) high with a matrix of 120 x 27 LED pixels. Each LED pixel matrix is a 9 x 5 display module and has a cluster of five red and three green closely spaced, discrete LEDs. The typical font employed in the RI VMSs is seven pixels tall by six pixels wide and has a two-pixel stroke width with two pixels inter-character spacing. The Vanguard® sign is able to display a three-line, 7 x 6 double stroke text message with a maximum of twenty 18-inch characters for each line. If displayed in one line only, it can have a character font of up to 72 inches. In addition, several dynamic features, including flashing, scrolling, and multiple-frame, are available for presenting messages. The VMS system also has graphics capabilities and is able to display symbols, graphic images, and icons.

Although several past studies have been conducted by Wang and Cao (*2, 3*) to study effects and impacts of portable variable message sign (PVMS) in RI, no systematic study has been done on these newly installed permanently overhead mounted VMSs. Since most of these systems are installed on arterial highways, it is important to know whether sign messages presented on these systems could be comprehended correctly and promptly by drivers moving at highway speeds, especially in high-volume traffic and construction/repair zones. In addition, it is critical to explore the impacts of these dynamic messaging features on drivers' responses to and comprehension of the messages.

As Rhode Island is a diverse society, the issues stated above might have a bigger impact to ethnic groups using other languages than English as their primary language. According to US Census 2000 (*4*), foreign-born residents accounted for 11.4 percent of the overall population in Rhode Island. This represents an increase of 25.4 percent from the 1990 foreign-born population and 53.9 percent of the state's overall population change. The continued growth in the foreign-born populations is an important factor that needs to be taken into account in this and other studies.

To this end, a human factors study was designed to investigate various features in VMS messaging and their impacts on a driver's understanding of and reaction to a message. The features considered here include: color schemes, number of message frames, flashing effect, wordings, and formats. Also considered here are demographic differences among drivers' age, gender, and native language. Special attention was paid to the ESL population (people using English as their second language) since they might find sign reading challenging during driving. It is important to know whether these factors

- affect a driver's ability to properly understand a message,
- affect the amount of time it takes a driver to read and comprehend a message, and
- influence the importance that a driver places on a message.

In this study, three approaches - driver questionnaire surveys, lab experiments, and field studies were employed to examine the various VMS messaging features. Section 2 outlines the research objective and goal, section 3 summarizes studies found in literature, section 4 presents the research methodologies, section 5 gives the results and discussions, and section 6 concludes the study.

2. RESEARCH OBJECTIVE AND GOAL

This research is aimed to examine how various VMS messaging features might affect drivers' understanding of and response to the message while driving. The objectives are:

- Conduct literature review and surveys to identify potential VMS display and format features that might affect motorist comprehension of the messages.
- Design and carry out lab simulation experiments to measure motorists' responses to various combinations of these features.
- Perform field studies on selected routes and correlate results with those obtained from lab simulation experiments.
- Analyze results and make recommendations and guidelines to RI highway authorities.

The overall goal is to help traffic engineers and highway management to design VMS messages that could enhance driving smoothness and safety on RI highways.

3. LITERATURE REVIEW

Survey results in the 2002 Intelligent Transportation Systems (ITS) Deployment Tracking database indicated that there are 2,805 VMSs installed in 97 major metropolitan areas in the United States (*5*). The data pointed out the fact that VMSs are widely implemented tools that serve as one of the main ITS components to promote road safety and improve traffic flow via effective communication to drivers. As highway authorities are increasingly relying on VMS to disseminate traffic information to drivers on highways, it is very important to understand how the messages influence drivers and how drivers respond to them. Some past studies related to these issues are briefed below.

3.1 VMS Impacts and Effects on Drivers

By surveying more than 500 drivers in the Washington, D.C. area, it was found that about half of the participants often responded to VMSs while 38% occasionally responded to VMSs (*6*). A stated-preference survey conducted in Wisconsin found that about 62% of the participants responded to arterial VMSs more than once per week and 66% of them changed their route at least once per month (*7*). An empirical analysis based on an extensive survey conducted in Amsterdam revealed over 70% of the drivers was sometimes influenced by VMS information (*8*). Another study conducted in Paris also found that 70% of drivers believe that VMSs are useful (*9*). A questionnaire survey concerning drivers' attitudes indicated that most of the drivers considered that VMS information can be very useful to them (*7, 10*).

When investigating the impacts on drivers based on information provided by VMS, it concluded that the impacts mainly depend on message content and message format (*11, 12, 13, 14, 15*). These analyses indicated that the content in terms of detail level of relevant information significantly affects drivers' willingness to make a diversion. In addition, driver behavior can be changed with appropriate travel information presented in a message. Thus, the more specific and clearer a message is, the more persuasive and influencing it becomes. As to message format, drivers' preference of traveling on a route could be affected by the format of the message though with identical information.

Many studies found that there were demographic effects on drivers' attitudes about and responses to VMSs (*2, 3, 8, 11, 16, 17, 18*). They found that women are less influenced by VMS information during the trip. Young people are less inclined to comply with VMS advice. Results from several lab experiments showed age effects existed across all tested conditions. Specifically, older subjects demanded longer response times but responded with less accuracy in most test conditions. In addition, questionnaires conducted by Nsour (*19*) showed that the task of reading VMS messages was one of the most difficult tasks for elderly as compared to young drivers. About 25% of the elderly surveyed viewed reading VMSs as either difficult or very difficult.

Although the main objectives of VMS are to enhance traffic flow and road safety, studies did show evidence of slow-downs when drivers approached the VMS messages (*15, 20*). In addition, a study of the effect of VMS on vehicle speeds and hourly vehicle speed deviations at the hourly level in Washington indicated that there was a significant decrease in mean speed when the VMSs were on, along with a significant increase in speed deviation (*21*). These indicated that drivers responded to VMS but require extra time to process the information provided by VMS. Further, an important finding indicated that drivers tended to compensate for the speed reduction from reading the messages by driving faster downstream after the slow-downs, which could result in safety concerns.

As stated above, many studies revealed that drivers constantly obtain information from VMSs in their daily driving experience. However, studies also showed that some drivers do not pay attention to VMSs. Chatterjee et al (*10*) employed questionnaire surveys to investigate the attitudes of drivers to VMS messages and the effect of different VMS messages on route choice. The results indicated that only one-third of drivers saw the messages presented to them and few of these drivers responded to the diversion suggestion. Another study conducted by Harder et al (*20*) with the application of a fully-interactive, PC-based STISIM driving simulator also found that about half of the participants did not response to the message as expected in the experiments. Three main explanations were given: first, they ignored the VMS message because they did not think that it applied to them (35.9%); second, they did not understand the VMS message (35.9%); and third, they did not notice the message (22.5%).

3.2 Design and Display of VMS Messages

The effectiveness of a VMS primarily depends on its message design and display format (*2, 3, 17*). Effective message design could help drivers understand the messages and make timely responses. General factors considered in message design include: color scheme, font size, number of message lines, wording, abbreviations, etc. Wang and Cao (*2, 3*) suggested that green and 5” x 7” were the best font color and font size for portable variable message signs (PVMSs). In addition, they found that messages with fewer lines (words) were responded to faster. Armstrong and Upchurch (*22*) carried out a field study in Phoenix, Arizona and concluded that the number of words in a message should vary with the VMS technology, the lighting conditions, and the prevailing traffic speed. VMS using different character fonts or dimensions should undergo a legibility analysis prior to implementation. Studies showed that specific, concise wordings of VMS messages would have a positive influence on drivers (*11, 13*).

To assess drivers’ understanding of abbreviations on VMS messages, two similar studies were conducted in Texas and New Jersey where participants were given a list of abbreviations and were asked to interpret the full words/phrases (*23, 24*). The results identified 24 abbreviations that were understood at an acceptable level for use on VMS, but regional differences with respect to driver understanding of some of the abbreviations were noted.

Two dynamic features of VMS, message alternating and message flashing, have been investigated extensively. Dudek and Ullman (*25, 26*) found that reading times were higher with flashing messages and suggested: one-frame VMS messages should not be flashed.; a line on a one-frame message should not be flashed; and a line on a two-frame message should not be alternated while keeping other lines the same. Simulation results also showed that static (one frame) messages took less response time than alternating (two frame) messages (*2, 3, 16, 17*). When a two-frame message has to be displayed, controlled field studies conducted by Dudek et al (*27*) showed that messages with better recall were displayed at 2 sec/frame (81%) and 4 sec/frame (85%). The common practice now is to display each frame for two seconds such that drivers could see a two-frame message displayed twice within the viewing distance. However, one study recommended that two-frame messages should be displayed at least 3 sec/frame to accommodate older drivers (*28*). No more than two frames should be displayed, as recommended by studies and government regulation (*23, 24, 29, 30*).

3.3 Three General Approaches to Study VMS

From the review of past studies given above, it noted that three general approaches were commonly used to investigate the effects and impacts of VMS systems. They were: driver questionnaire surveys, lab simulation experiments, and field studies. Driver questionnaire surveys collect respondents' opinion or choice through a list of multiple choices or open-ended questions. No real-time driving performance data are collected. Results from one site are not always transferable to other sites due to differences in network characteristics and attitudes and experiences of drivers. However, a questionnaire survey is still considered as a cheaper, reasonable method that could be quickly adapted to a new study and yield valuable results (*10, 12*). Specifically, the surveys employing stated-preference technique may be able to provide more accurate results, since participants are normally asked to describe how they actually behaved under given conditions (*7*).

Driving simulation measures drivers' responses to artificially introduced VMS stimuli in a simulated driving environment. Participating drivers, sitting behind a steering wheel of a stationary test vehicle, usually experience some degree of virtual driving in a lab setting with no risks. By its nature, lab simulation allows researchers a lot more freedom to experiment with nearly every possible variation of VMS. As variables examined in the experiment are under strict control, the experiments are mostly repeatable and results can be thoroughly analyzed by statistical methods. However, the interpretation of results is sometimes questioned since test drivers are not encumbered by actual driving in the simulation experiment (*10, 12, 31, 32*). Further, the sample sizes of subjects used in the highly structured lab experiments are normally small and typically fewer than 30 subjects (*33*). To address these questions, studies have been carried out to validate the results from the laboratory-based driving simulation. By comparing route choices in real life with those based on VLADIMIR route choice simulator, Bonsall et al (*34*) concluded that a well designed simulator is able to precisely replicate the route choices with a very high degree of detail and accuracy. In addition, evidence suggested that it can accurately replicate route choice responses to roadside VMS information. Another study conducted by Lee et al (*35*) compared simulated driving performance with actual on-road driving performance among older drivers. They found high positive relationship between simulated driving and on-road assessment and suggested that driving simulators can serve as a safe and economical means of assessing driving performance on elder drivers.

Field studies, on the other hand, provide actual results regarding drivers' response to VMS in real driving conditions. Study results can be seen to apply to real traffic environments immediately. However, field experiments are usually confined by the limitation of actual in-service VMS on the test route. These experiments are usually time-consuming, more risky and stressful to participants, especially to older drivers (*36*, *37*). Experiments are not repeatable due to many uncontrolled variables, such as dynamic traffic and weather conditions in the complex real driving environment (*32*). In addition, liability issues involved in the field study could pose difficulty to the investigators and participants.

4. METHODOLOGY

Three approaches were employed in this study to examine the effects of certain factors and their combinations on drivers' comprehension of and response to VMS messages. The three approaches are: questionnaire surveys, lab experiments, and field studies. Questionnaire surveys collect drivers' opinions and preference through a number of multiple-choice questions. Lab experiments measure drivers' responses to a series of VMS images in a simulated driving environment. Field studies capture drivers' responses to real in-service VMS during actual driving. The three methodological approaches are described in more details below after a description about subjects participated in the study.

4.1 Subjects

To ensure the rights and welfare of the human subjects involved in this research, researchers were certified and the research plan was approved by the Institutional Review Board (IRB) of the University of Rhode Island prior to commencing the study.

Thirty-six subjects were recruited to participate in the survey and lab experiment, while eighteen of them voluntarily participated in the field study. Participating subjects were required to possess a valid RI driver license, driving experience on interstate highways, and normal or near-normal eyesight. To balance the age and gender effects, equal numbers of males and females were recruited from three age groups: 20-40 years old, 41-60 years old, and 61 and above.

Each subject was asked to sign a consent form and fill out a demographic questionnaire prior to starting their tasks. Subjects were then briefed and instructed about the purpose and procedures and told that they could end participation at anytime during the study. It was assured to each individual subject that all records would be kept confidential and only used for the purpose of the present research. Subjects were informed that they would be compensated for their participation in the survey and lab experiment, and those who volunteered to participate in the field study would receive additional stipend.

4.2 Questionnaire Survey

A questionnaire was developed to collect drivers' opinions and preferences regarding the display of VMS messages. A total of forty-four survey questions were presented to subjects (see attached CD.) Each question presented multiple choices of VMS messages shown in different color schemes, formats, or wordings. Microsoft PowerPoint® was employed in the development and presentation of the survey question since it is capable of replicating all of the dynamic features seen on actual VMSs. The messages were selected from a library approved by the Rhode Island Department of Transportation (RIDOT). They represented common highway scenarios in Rhode Island such as: accidents, congestion, construction, and icy roads. The selected messages are shown in Table 1. Each question in the survey was presented with multiple choices of VMS images, each displaying identical or similar content on a standard black background. These multiple VMS images within a question represented variations of one messaging factor that might affect drivers' responses. Only one factor was surveyed in a question at a time. This technique allowed a systematic analysis to help determine the most preferred effect regarding a certain messaging factor.

Table 1. VMS messages employed in the questionnaire survey

Message number	Message content
1	ACCIDENT AT EXIT 12 MAJOR DELAYS TO BOSTON USE ROUTE I-295
2	CONGESTION 1 MILE AHEAD MINOR DELAYS TUNE RADIO TO 1610AM
3	CONSTRUCTION AHEAD LEFT LANE CLOSED KEEP RIGHT
4	ICY ROAD BETWEEN EXITS 10 AND 15 REDUCE SPEED
5	ACCIDENT 1 MILE AHEAD USE ROUTE I-295

Table 2 lists the factors considered in the survey and their respective effects. Specifically, these factors consisted of: number of frames (one-frame and two-frame), flashing (static, all flashing, and one-line flashing), color (red, green, and amber), color combinations (all amber, green & amber, red & amber, and tri-color), wording (less specific, somewhat specific and very specific), and abbreviations (no, little, some, and many). The selected factors and effects are within the capabilities of the VMS system currently used in Rhode Island.

Table 2. Factors and effects considered in the questionnaire survey

Factor	Factor Effect
Number of frames	One-frame, Two-frame
Flashing	Static, All flashing, One-line flashing
Color	Red, Green, Amber
Color combinations	All amber, Green & amber, Red & amber, Tricolor
Wording	Less specific, Somewhat specific, Very specific
Abbreviations	No, Little, Some, Many

As mentioned, different choices within a question shared the same message content but were displayed differently with respect to a factor. Each question investigated only one factor at a time. Each factor was tested four times using different messages. The sequence of survey questions and the choices presented within each question were randomized to prevent any biased response. A detailed description about each factor and its respective effects follows.

- Number of Frames: Three choices of identical messages were presented in either one static frame or two alternating frames. The two alternating frame messages were displayed with two different frames alternating at 2 sec/frame. An example is given in Figure 1.

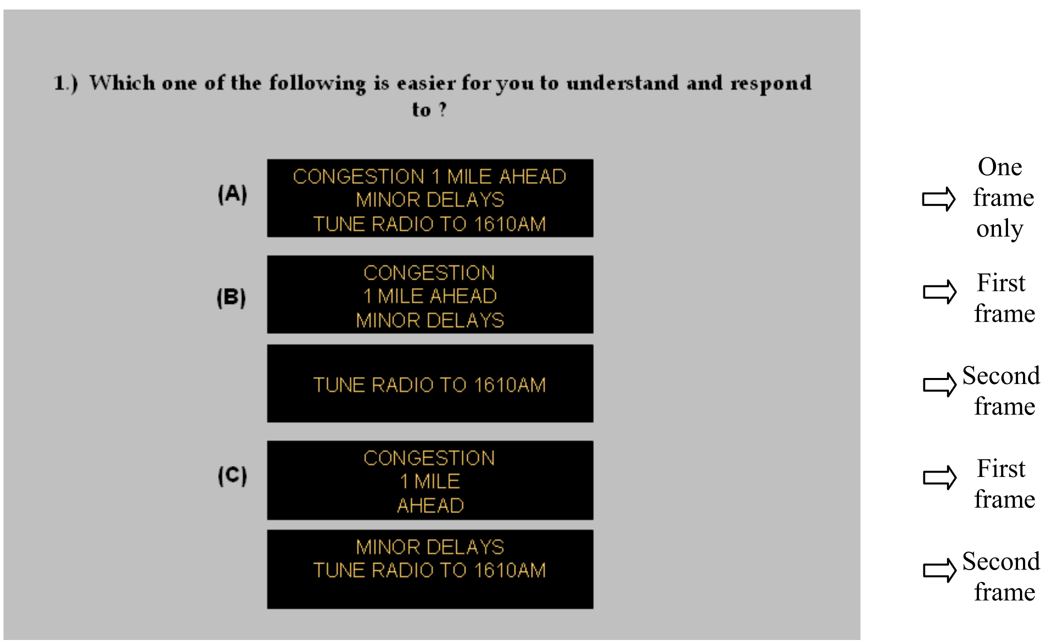

Figure 1. Sample survey question to test the number of frames effect

- Flashing: Three choices of identical messages with varying flashing effect were presented, one with a one-line flashing message, one with all three lines flashing, and one with no flashing. The flashing line(s) was flashed twice per second. An example is shown in Figure 2.

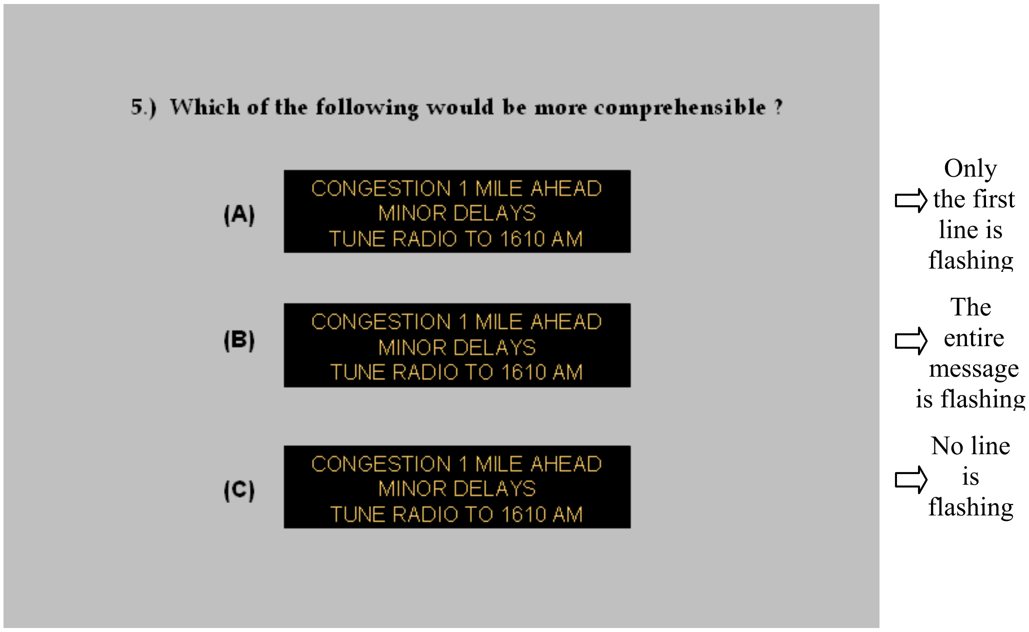

Figure 2. Sample survey question to test the flashing effect

- Color: Three identical messages in different colors were presented, one in red, one in green, and one in amber. An example is shown in Figure 3.

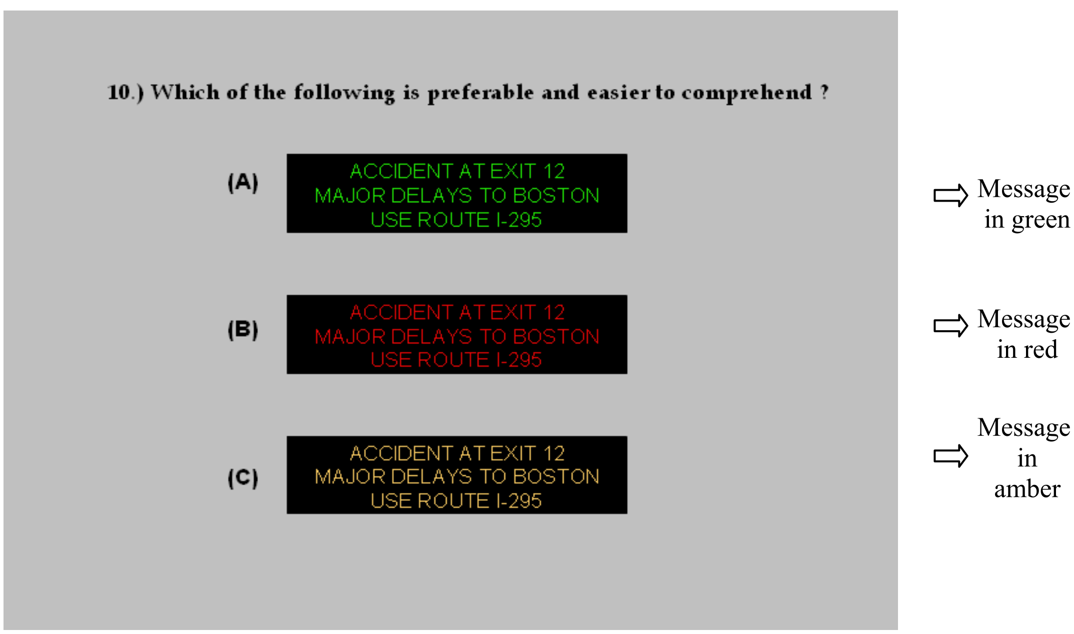

Figure 3. Sample survey question to test the message color effect

- Color Combinations: Four choices of identical message with various color combination were presented, one with all amber lines, one with the first line in green and the second and third lines in amber, one with the first line in red and the second and third lines in amber, and one with the first line in red, the second line in amber, and the third line in green. An example is shown in Figure 4.

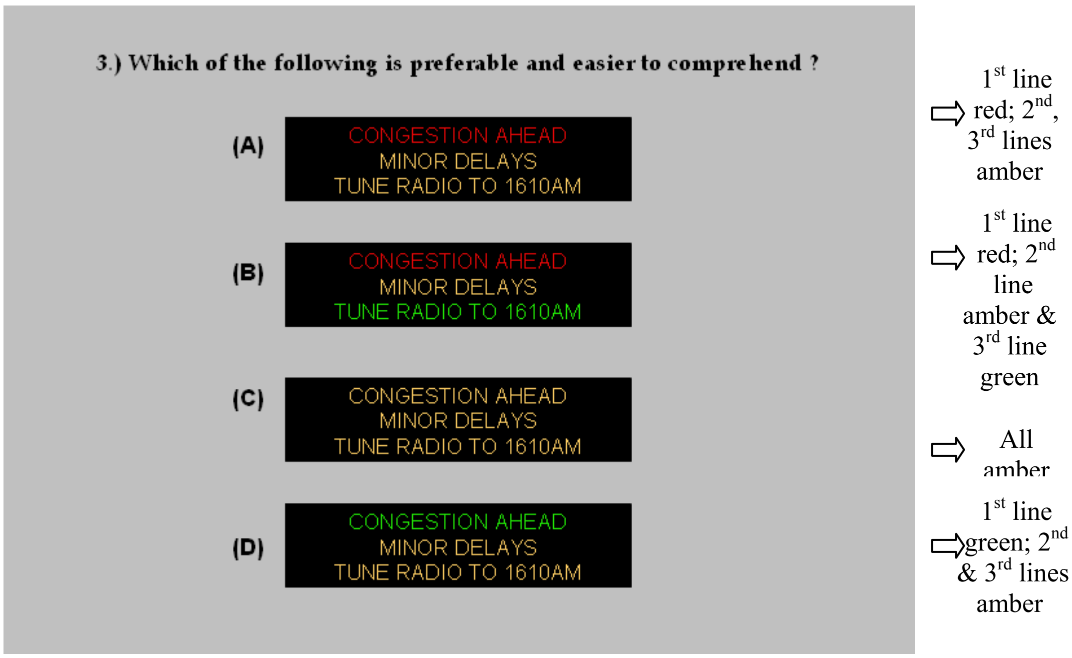

Figure 4. Sample survey question to test message color combination effect

- Wording: Three similar messages varied in their wordings were presented as choices, one less specific, one somewhat specific, and one very specific. An example is given in Figure 5.

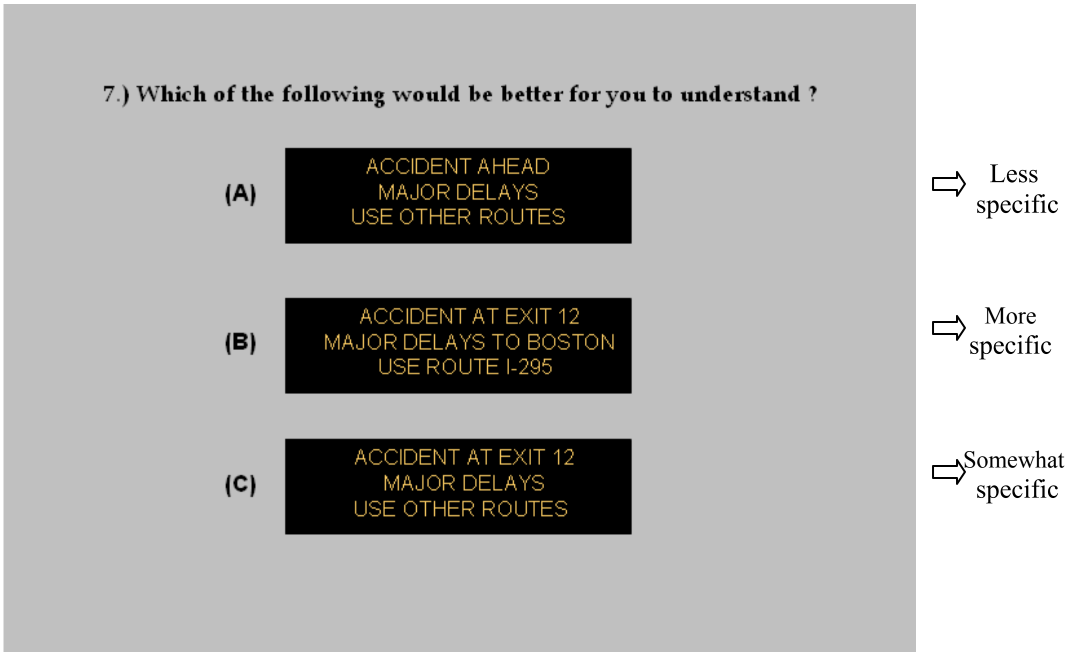

Figure 5. Sample survey question to test the wording effect

- Abbreviations: Four similar messages varied in their use of abbreviation were presented as choices, one without abbreviation, one with little abbreviations, one with some abbreviations, and one with many abbreviations. An example is shown in Figure 6.

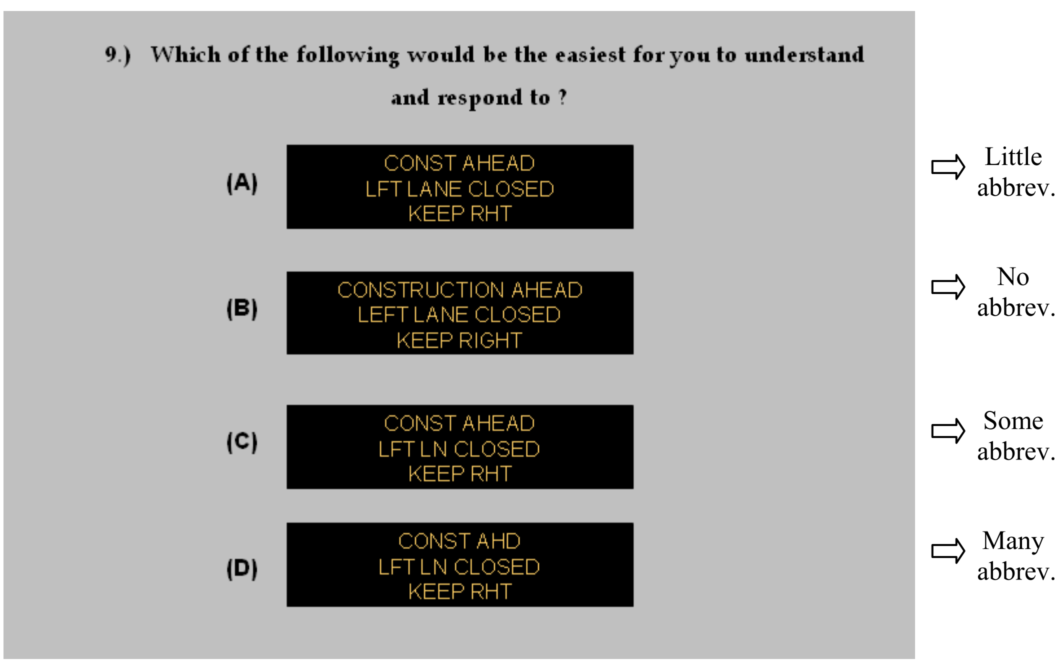

Figure 6. Sample survey question to test the abbreviation effect

When a subject started taking the survey, she/he would first see an instruction slide (see Figure 7), followed by another slide regarding the subject's demographic information (see Figure 8.) Survey questions were then presented one at a time. The subject was asked to choose an image that she/he most preferred and then mark the answer on the answer sheet. After all forty-four questions were completed, the subject was solicited for feedback and comments. Feedback and comments collected from subjects were documented for reference.

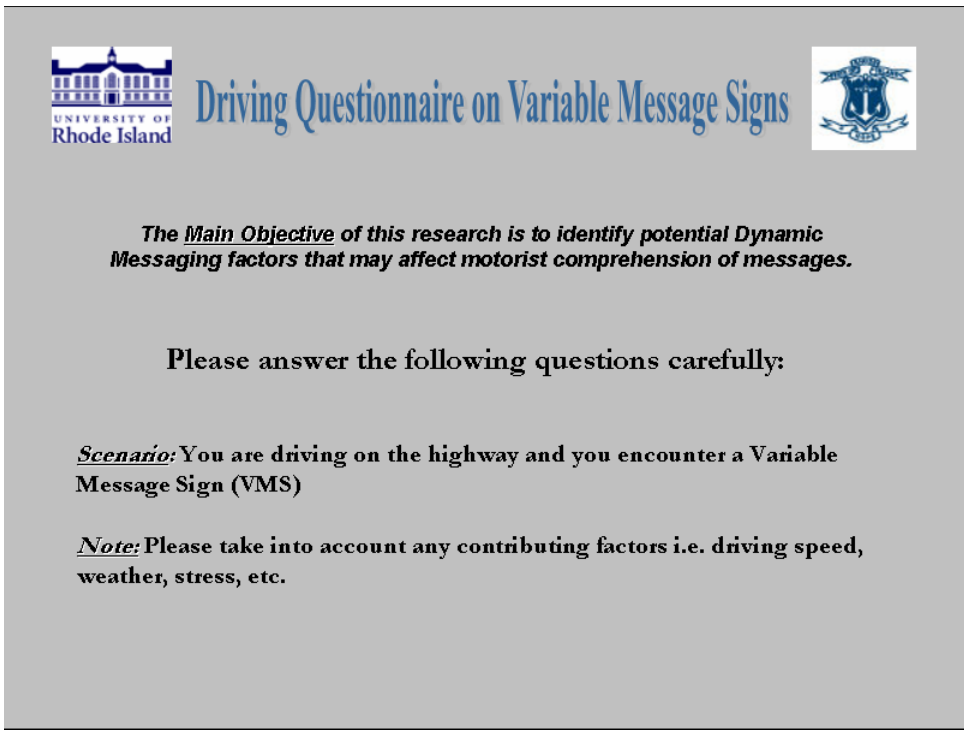

Figure 7. Instruction slide for the questionnaire survey

Please provide the following demographic information for analysis purposes. (Place an X on the squares to give your answer)

Name: ____________________ E-mail: ____________________

Contact Address: ____

Phone Number: ____________

Gender: ☐ Male ☐ Female Corrective Lenses: ☐ Yes ☐ No

Education (select the highest that you have completed): ☐ elementary

☐ high school ☐ 2-yr college ☐ 4-yr college ☐ post college

Age Group: ☐ 20 – 40 yrs ☐ 41- 60 yrs ☐ > 60yrs

Highway Experience: ☐ Less than 1yr ☐ 1 – 3yrs ☐ 4 -6 yrs ☐ 7 or more

Native Language: ☐ English ☐ Spanish ☐ Other ____________
(please specify)

Figure 8. Demographic information request for questionnaire

4.3 Lab Experiment

Lab experiments were set up to measure subjects' response times to various VMS messages displayed in a simulated driving environment. Human-factors and design of experiment principles were carefully followed in the design and development of the lab experiments. Two groups of factors, within-subject factors and between-subject factors, were considered in the experiment as exhibited in Table 3. Within-subject factors included flashing effect and color combination, while between-subject factors were age and gender. These factors were key factors in this study since they provided a means to gauge driver's responses to various VMS messages and a means to contrast survey findings with lab experiment results.

Table 3. Experiment factors and their levels

Within-subject Factors	Level
Flashing effect (F) Color combination (C)	Static (1), All flashing (2), Line flashing (3) Amber (1), Green & amber (2), Red & amber (3), Tricolor (4)
Between-subject Factors	
Subject's age (A) Subject's gender (G)	20 ~ 40 (1), 41 ~ 60 (2), Above 60 (3) years old Female (1), Male (2)

Considering age and gender as blocking factors, a blocked factorial experiment design was employed in the study. The statistical model was:

$$T = \mu + F_i + C_j + (FC)_{ij} + A_k + G_l + (AG)_{kl} + \varepsilon_{ijkl} \quad (4.1)$$

where:

T – subject's response time (in seconds);

μ – overall mean;

F – flashing effect (i = 1, 2, 3);

C – color combination (j = 1, 2, 3, 4);

(FC) – interaction effect between F and C;

A – subject's age (k = 1, 2, 3);

G – subject's gender (l = 1, 2);

(AG) – interaction effect between A and G;

ε – error.

Three sets of VMS messages were employed in the experiment (see Table 4.)

Table 4. Three VMS messages used in the experiment

Message 1	Message 2	Message 3
ACCIDENT 1 MILE AHEAD USE ROUTE I-295	SEATBELTS SAVE LIVE CLICK IT OR TICKET	TRAFFIC INFORMATION TUNE TO 1610 AM

All VMS images were generated by the Vanguard® VMS Central Controller, the same software used by RIDOT. Full matrix text messages were displayed on a typical black background (Figure 9.) The VMS image generated was to mimic the Daktronics Vanguard® VMS system (model VF-2000-27x120-18-W) that is currently in-service in Rhode Island. Since the dynamic features could not be shown on paper, a detailed description regarding the flashing effect and color combination on VMS images seen in the experiment are given in Table 5.

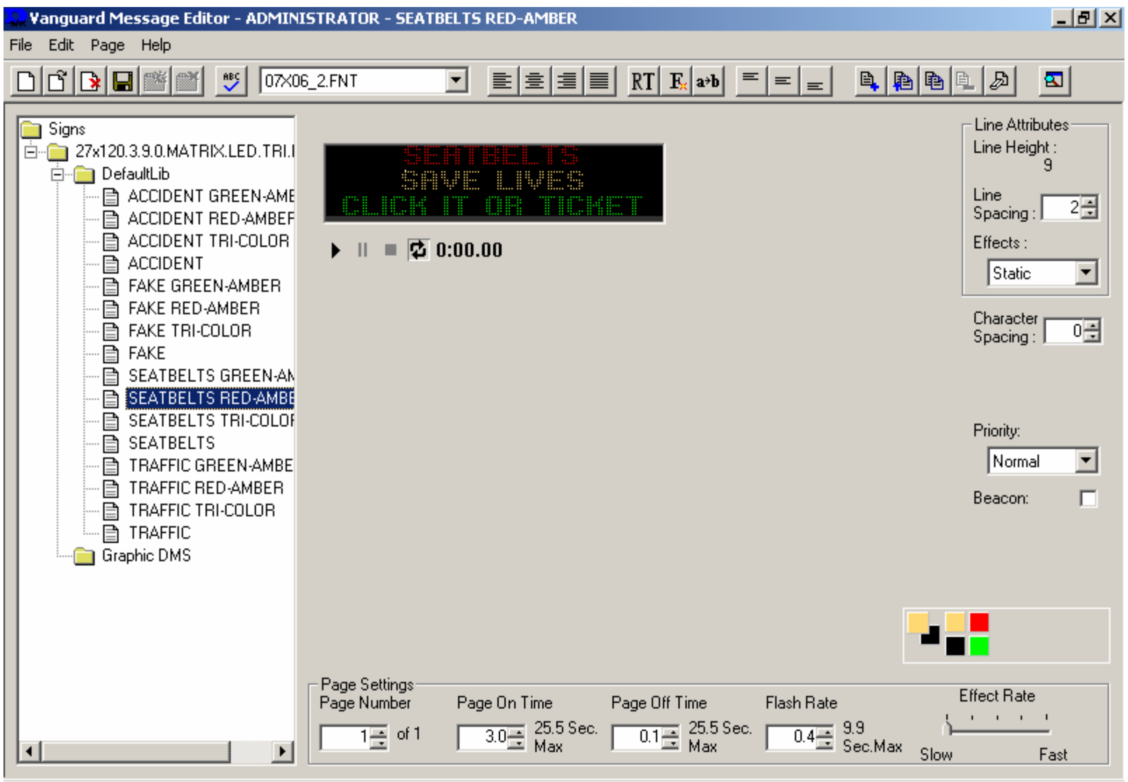

Figure 9. A screen shot of the Vanguard Message Editor

Some fake messages were added to the experiment. This was to prevent a subject from making a response too early based on the first word of a message. These fake messages imitated real test messages but consisted of inappropriate combinations of words chosen from the real ones (see Table 6.) A total of twelve fake messages were included in the experiment. With the addition of fake messages, a subject would need to read through all three lines of messages before she/he could correctly identify the message. The 3 test messages were displayed with 3 flashing effects and 4 color combintions. With two replications on all test messages (including fake), a total of 96 VMS messages were presented to the subject in a random sequence during the experiment.

Table 5 Flashing effect and color combination applied to VMS images

		Color Combination			
		All amber	Green & amber	Red & amber	Tricolor
Flashing	Static	● Message was in amber only. ● Message was static.	● First line of message was green; the other two lines were amber. ● Message was static.	● First line of message was red; the other two lines were amber. ● Message was static.	● First line of message was red, the second line was amber, third line was green. ● Message was static.
	All flashing	● Message was in amber only. ● Entire message was flashing.	● First line of message was green; the other two lines were amber. ● Entire message was flashing.	● First line of message was red; the other two lines were amber. ● Entire message was flashing.	● First line of message was red, the second line was amber, third line was green. ● Entire message was flashing.
	Line flashing	● Message was in amber only. ● First line of the message was flashing; the other two were static.	● First line of message was green; the other two lines were amber. ● First line of the message was flashing; the other two were static.	● First line of message was red; the other two lines were amber. ● First line of the message was flashing; the other two were static.	● First line of message was red, the second line was amber, third line was green. ● First line of the message was flashing; the other two were static.

A real video based simulation was employed in the lab experiment. The simulation video clips were created by merging real driving videos with digitally created VMS images. The driving video was taken in the afternoon of a mid-March day in 2004. Driving speed was kept constant at 50 mph while driving from exit 6 to exit 5 on RI Route 4 heading southbound. This

Table 6. Example fake VMS messages used in the experiment

Fake Message 1	Fake Message 2	Fake Message 3
TRAFFIC 1 MILE AHEAD CLICK IT OR TICKET	SEATBELTS SAVE LIVES TUNE TO 1610 AM	TRAFFIC 1 MILE AHEAD USE ROUTE I-295
Fake Message 4	Fake Message 5	Fake Message 6
ACCIDENT SAVE LIVES CLICK IT OR TICKET	SEATBELTS INFORMATION TUNE TO 1610 AM	SEATBELTS INFORMATION USE ROUTE I-295

route was chosen because of its light traffic and the presence of an in-service VMS in between the two exits. A Canon XL1 digital video camcorder, mounted on a tripod inside a 2001 Chrysler Voyager and leveled at driver's eye height, was used to take the video. The digital video was later downloaded onto a desktop computer through a firewire cable where individual frames were extracted by Sonic Foundry VideoFactory™. Individual VMS image was created, resized, copied, and then pasted onto the VMS board in each frame of the driving video. Consecutive frames were next rendered into a video clip at a rate of fifteen frames per second. Each rendered video clip lasted 27 seconds and was stored in NTSC DV avi format at 720 x 480 pixel resolution with a refresh rate of 29.970 (fps).

The setup of this experiment is shown in Figure 10. The main elements of the experimental apparatus include:

- A four-door 1998 Ford Taurus sedan – a full-size, stationary vehicle to accommodate test subjects in the experiment,
- A Dell Dimension 4500 server with an enhanced video processor – a multi-functional computer used in the study to administer the lab experiment and to record experiment data,

- A Microsoft Sidewinder force feedback wheel – an interface device between a test subject and the computer via a USB connection to capture a subject's response,
- A BenQ PB8230 DLP digital projector with 1024 x 768 XGA resolution and 2500 ANSI Lumens – an overhead mounted projector to project the driving simulation video onto a wide screen, and
- A Draper's Cinefold projection screen (12 feet wide x 6 feet and 9 inches high) – a front projection, flat-surface, tensioned, wide screen with 16 x 9 aspect ratio that is used for simulation video projection.

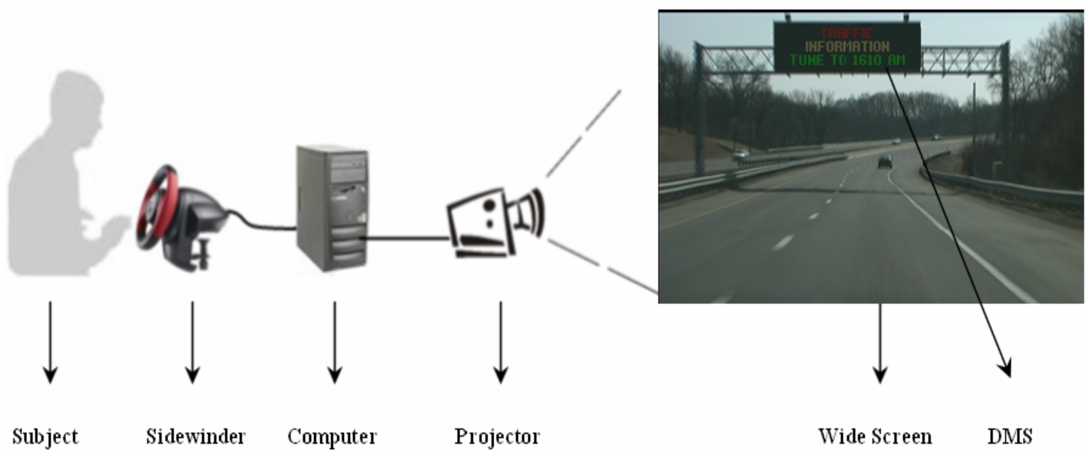

Figure 10. A schematic diagram for the experiment

The front bumper of the Taurus was 4 feet away from the wide screen. The distance between the screen and the subject sitting in the driver's seat of the vehicle was 7 feet and 4 inches. The distance between the subject's eyes and ground surface was approximately 3 feet and 6 inches. The projector was mounted at a height of 7 feet and 4 inches from the ground surface at a distance of 18 feet and 6 inches from the wide screen.

Prior to the start of experiment, a research assistant would give the test subject a briefing about the simulation experiment, answer any questions, and ask the subject to read and sign a consent form. The test subject would next be led to the lab where all lights were shut off except the interior light of the test vehicle. The subject would be asked to sit in the driver's seat of the vehicle and make herself/himself comfortable by adjusting seat position, seat height, vent, and lights. The subject was instructed to press one of the four pre-defined buttons in the Sidewinder wheel according to the content of the VMS message regardless of its color combination or flashing effect. She/he would press button "1" for message 1, "2" was for message 2, "3" was for message 3, and "4" for a fake message (see Figure 11.) An instruction sheet was also placed on the instrumentation panel in the test vehicle to further assist subjects with response button selection. It was stressed that both response speed and accuracy were important in the experiment. Specifically, they were told that fast but inaccurate responses was not desired. Additionally, subjects were alerted that they would encounter messages that were worded inappropriately and should consider them as fake messages. To encourage response accuracy, subjects were told that they would earn a special gift in addition to the stipend if they correctly identified 90% or more of the VMS messages (see Figure 12.) These fake messages would require them to make sure each image was read in their entirety for coherency so as to preserve accuracy ratings.

Figure 11. Key pressing instructions for different VMS messages

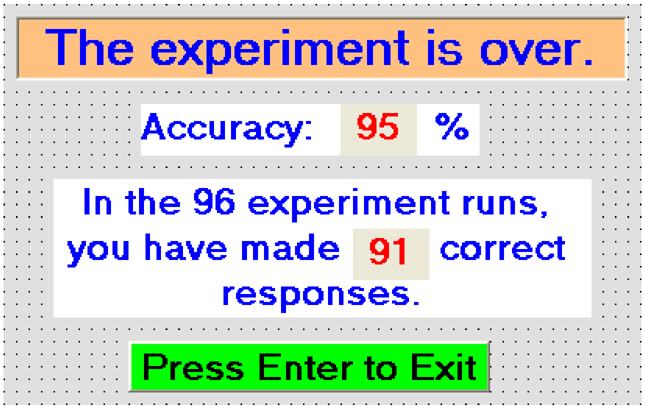

Figure 12. Accuracy results

Some past studies showed that motion or simulation sickness was occasionally reported by subjects participating in driving simulation experiments, especially by elder subjects (*35, 38, 39*). A short practice run was given at the beginning of the experiment to help identify subjects who were more vulnerable to this type of problem. If identified with motion sickness, the subject could determine whether to continue with the actual experiment or not. Another function of the practice run was to familiarize the subject with this simulation experiment to help her/him reach consistent performance. The subject could repeat the practice run if she/he desired. After the practice run, the subject was asked to enter some basic demographic information and signal the research assistant to start the experiment (see Figure 13.) During the lab experiment, video clips containing digitally generated VMS images were projected onto the wide screen in front of the Taurus (see Figures 14 & 15.) In each clip of the simulation video, the VMS image would initially appear as a small dot and gradually increase in size as seen in actual driving. Ninety-six video clips with various VMS images were shown to a subject in a random but controlled manner administered by a computer server. The Sidewinder, a PC-governed feedback device replacing the original steering wheel, was used to capture subjects' responses throughout the experiments. Subjects would respond to the VMS stimuli by pressing the pre-defined buttons on the Sidewinder, as instructed previously. The experiment had two sections with a break in between. The length of the break was controlled by the subject (see Figure 16.) During the experiment, the subject could be alerted if no response was detected before passing the VMS sign (see Figure 17) or if a response was made prior to the appearance of the VMS sign (see Figure 18.)

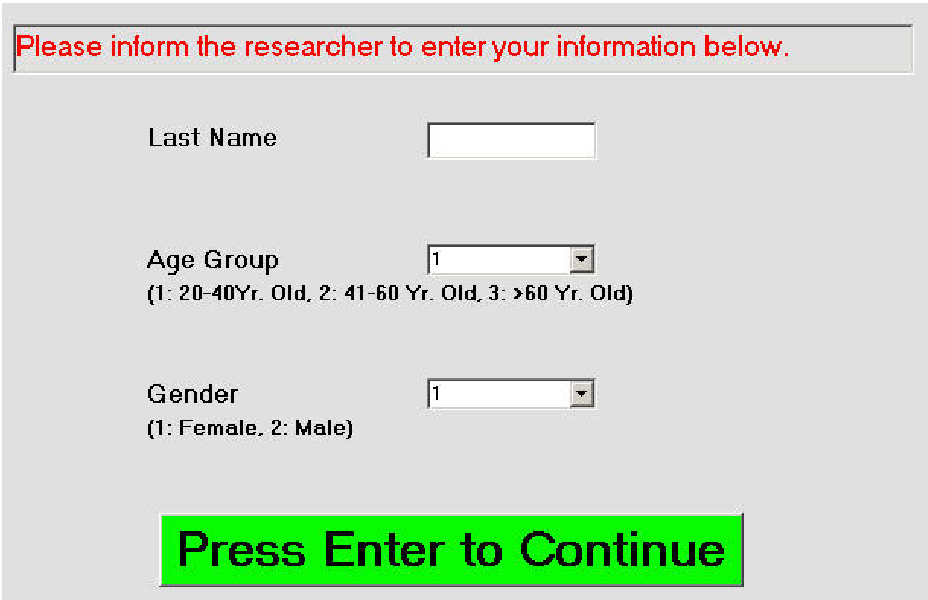

Figure 13. Subject information form

Figure 14. A screen shot

Figure 15. A shot of the experiment in progress

Figure 16. Message box for taking a break

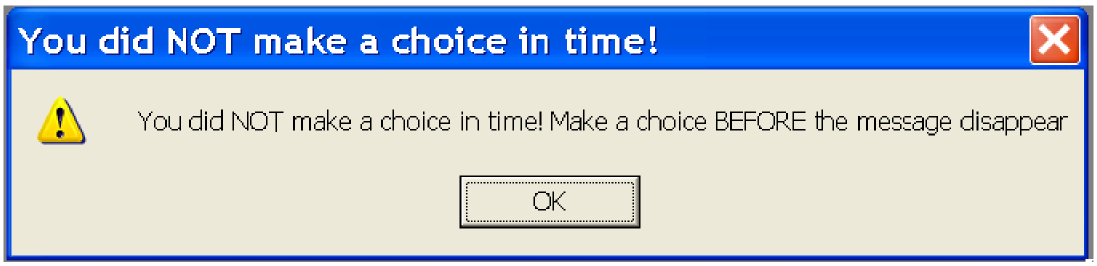

Figure 17. Alert box for not making response

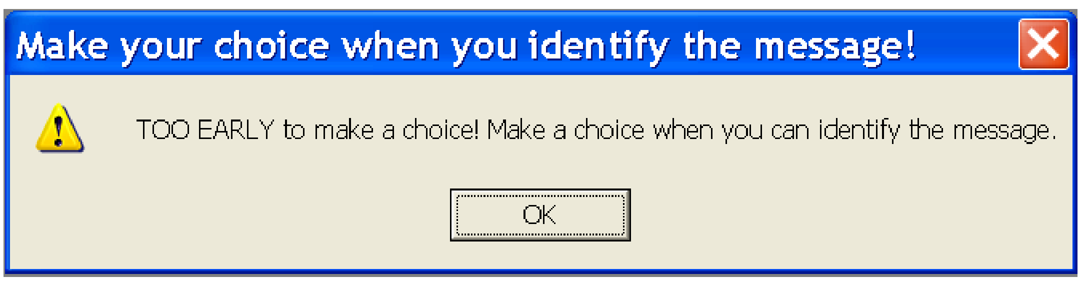

Figure 18. Alert box for making response too early

A subject's response time and accuracy to each VMS image were recorded in a Microsoft Access® database in the computer server. Response time and accuracy were two performance measures used in this lab experiment. Response time was measured as the time difference between the starting of the simulation video clip and the subject's response to the VMS stimulus. An accuracy statistic was calculated for each subject as the ratio of the correct responses made to the total number of stimuli. Each record in the database file corresponded to a clip of VMS simulation. It recorded the name, age, gender of the subject; the content, color combination, flashing effect, response time, response key, response correctness of the response (see experiment data example in Appendix A.) The overall experiment including a break took about 40 to 50 minutes to complete.

4.4 Field Study

To better understand a driver's comprehension of a VMS message in actual driving, a field study was designed. Subjects participating in the lab experiment were invited to participate in the field study on a voluntary basis. The study was done on the same highway segment where the driving video was taken. The main objective here was to explore the correlations between the field study results and the lab experiment results on a subject-by-subject basis. A strong correlation would help develop a model that might be useful to predict driver response to VMS messages in real driving via simulated lab experiments. Only the "SEATBELTS" message was used (see Figure 19), since other messages might distract divers and were not appropriate to be displayed. The message displayed was a static, one-framed, tri-color message.

Figure 19. VMS message used in the field study on Route 4

Prior to starting the field study, the subject was briefed with the purpose and procedure involved. The test driver was asked to drive her/his own vehicle in the right lane on the test route and maintain a speed of 50 mph while driving through the test route. She/he was asked to read aloud the message as soon as it became visible. An in-vehicle digital camcorder, held by a research assistant sitting on the passenger seat, was used to capture the driving scene including the gradual appearance of the VMS and the driver's vocal reiteration of the message. Video recording started after passing a reference point and stopped after passing the VMS sign. Two runs were conducted by each participant and they took less than an hour to complete. The response time to the VMS was determined through a frame-by-frame analysis of the digital video. Response time was measured as the time difference between the reference point and the subject's vocal response to the VMS. The average response time of the two trials was later compared with the average response time to the same VMS message in the lab experiment.

5. RESULTS AND DISCUSSIONS

5.1 Questionnaire Survey

The questionnaires collected participants' preferences about six VMS messaging factors including: number of frames, flashing, color, color combinations, wording, and abbreviations. The choices made by subjects were entered into a Microsoft Excel® worksheet where preference percentages were automatically calculated via macros. The results are shown in Table 7.

Table 7. Overall preference percentages on VMS effects from survey

Factor	Factor Level	Preference Percentage
Number of Frames	One Frame	61.84%
	Two Frame	38.16%
Flashing	Static	54.29%
	One Line Flashing	36.50%
	All Flashing	9.21%
Color Combinations	All amber	33.55%
	Green & Amber	38.16%

	Red & Amber	7.24%
	Tricolor	21.05%
Color	Amber	53.29%
	Green	44.18%
	Red	2.62%
Wording	Less Specific	21.05%
	Somewhat Specific	32.90%
	Very Specific	46.05%
Abbreviations	No Abbreviation	91.54%
	Little Abbreviations	3.85%
	Some Abbreviations	3.29%
	Many Abbreviations	1.32%

In addition, paired t-tests were conducted to test the significance between different variations with respect to each factor. It found that *one-frame* messages were significantly preferred over *two-frame* messages (p=0.027). From the post-survey comments collected from subjects, most of them thought "…two-frame messages are difficult to follow…" and "…need to wait to see the second frame for getting the whole information…".

Regarding the flashing effect, more than half of the subjects strongly preferred the *static* messages over the *one-line flashing* (p=0.021) or *all flashing* messages (p=0.000). When comparing the one-line flashing messages with the all flashing messages, most subjects favored the one-line flashing message over the all flashing message (p=0.000). Those who favored static messages thought "…flashing messages are annoying and disturbing…".

When VMS messages were displayed in solid colors, the majority preferred *amber* (53.29%) followed by *green* (44.18%) and *red* (2.62%). There was no significant difference between amber and green (p=0.443) but amber and green were significantly preferred over red (p=0.000).

When multiple colors were used to display different lines of the message, a *green and amber combination* (38.16%) was the most preferred; followed by *all amber* (33.55%) and *tricolor* (21.05%). No significant differences were found among these three. The *red and amber combination* (7.24%) was the least preferred and was significantly different from all amber (p=0.002), green & amber (p=0.000), and tricolor (p=0.045) displays. Most of the subjects thought "…messages containing red are difficult to read and identify…".

When messages were worded differently, most subjects strongly favored those with *very specific* wordings over those with either *somewhat specific* (p=0.018) or *less specific* wordings (p=0.000). They thought "…specific VMS messages are more informative…". When abbreviations were presented, an overwhelming majority (91.54%) indicated that they strongly preferred a message without any abbreviation whatsoever (p=0.000). Most of them thought "…it's not easy to understand a message with abbreviations…".

The preference percentages were also calculated by subject's demographics. The results are reported in Table 8. Statistical tests were also conducted to test whether there were any significance differences among individual groups with respect to the same demographic factor. It found that the preferences made by individual groups were mostly in-line with those made by the whole group. Some exceptions are reported below per individual demographic factor.

Table 8. Preferences percentages on VMS factor effects by demographics

		Gender		Age (yrs)			Language	
		Male	Female	20-40	41-60	61-above	English	ESL*
Frame	One Frame	73.61%	56.94%	50.00%	72.92%	72.92%	69.23%	55.00%
	Two Frame	26.39%	43.06%	50.00%	27.08%	27.08%	30.77%	45.00%
Flashing	Static	52.78%	61.11%	41.67%	66.67%	66.67%	58.65%	45.00%
	One Line Flashing	33.33%	34.72%	45.83%	29.17%	27.08%	35.58%	30.00%
	All Flashing	13.89%	4.17%	12.50%	4.17%	6.25%	5.77%	20.00%

Color Combo	All amber	37.50%	30.56%	27.08%	47.92%	27.09%	42.31%	12.50%
	Green & Amber	33.33%	44.44%	37.50%	39.58%	39.58%	33.65%	52.50%
	Red & Amber	5.56%	8.33%	6.25%	8.33%	6.25%	6.73%	7.50%
	Tricolor	23.61%	16.67%	29.17%	4.17%	27.08%	17.31%	27.50%
Color	Amber	52.78%	55.56%	56.25%	70.83%	35.42%	57.69%	45.00%
	Green	44.44%	41.67%	39.58%	27.08%	62.50%	40.38%	50.00%
	Red	2.78%	2.78%	4.17%	2.08%	2.08%	1.92%	5.00%
Wording	Less Specific	19.44%	20.83%	22.92%	10.42%	27.08%	22.12%	15.00%
	Somewhat Specific	33.33%	33.33%	29.17%	33.33%	37.50%	32.69%	35.00%
	Very Specific	47.22%	45.83%	47.92%	56.25%	35.42%	45.19%	50.00%
Abbreviations	No Abbrev.	93.06%	88.89%	89.58%	93.75%	89.58%	89.42%	95.00%
	Little Abbrev.	4.17%	4.17%	4.17%	4.17%	4.17%	4.81%	2.50%
	Some Abbrev.	1.39%	5.56%	4.17%	2.08%	4.17%	4.81%	0.00%
	Many Abbrev.	1.39%	1.39%	2.08%	0.00%	2.08%	0.96%	2.50%

* ESL: Those who use English as a second language.

Gender Factor. The only messaging factor that resulted in a somewhat different preference between males and females was the number of frames. Males strongly preferred one frame over two frames (p=0.033) while females did not. All other preference statistics were similar between the two genders.

Age Factor. When looked at the number of frames factor, it found that subjects over 40 years old significantly favored the one-frame messages over the two-frame, alternating messages while young subjects (20~40 yrs old) were indifferent between the two. With respect to the flashing effect, subjects over 40 strongly preferred static over flashing messages while young subjects were indifferent between static and one-line flashing messages. When a message was displayed in a solid color, amber was the most favored color by both young and middle-aged subjects while older subjects preferred green.

Native Language Factor. To find out whether a person' native language might have an impact on her/his preference toward a specific feature on VMS messages, survey results were compared between those subjects using English as their primary language (English) and those using English as a second language (ESL.) The ESL group had ten subjects while the English group had twenty-six. As Table 8 indicated, no significant difference was found between the two groups regarding number of frames, flashing effect, wording, and abbreviations. The ESL group preferred green & amber multi-colored VMS (p=0.032) while the English group preferred all-amber (p=0.000). When looking at solid-colored VMSs, the ESL group preferred green while the English preferred amber, although no strong disparity was found between the two groups (p=0.219 for amber; p=0.381 for green; p=1.000 for red). Regarding wording and abbreviations, it might be reasonable to presume that reading messages displayed in a language other than one's native language could be somewhat challenging. The survey results did confirm that the ESL group had a stronger preference for very specific, unabbreviated messages than the English group.

5.2 Lab Experiment

A total of 3,456 responses were collected from 36 subjects who participated in the lab experiments. With 215 inaccurate responses and 18 unusual observations removed, 3,223 observations were included in the statistical analysis. The data were analyzed by MINITAB® where analysis of variance (ANOVA) procedure was performed. The normal plot of residuals shown in Figure 20 indicates that the assumption of normality was upheld. An overall ANOVA result is summarized in Table 9. Flashing, color combinations and driver's age were significant at the 0.05 significance level. The interaction between flashing and color combinations and the interaction between age and gender were also significant. Main effect and interaction plots are shown in Figures 21, 22, and 23.

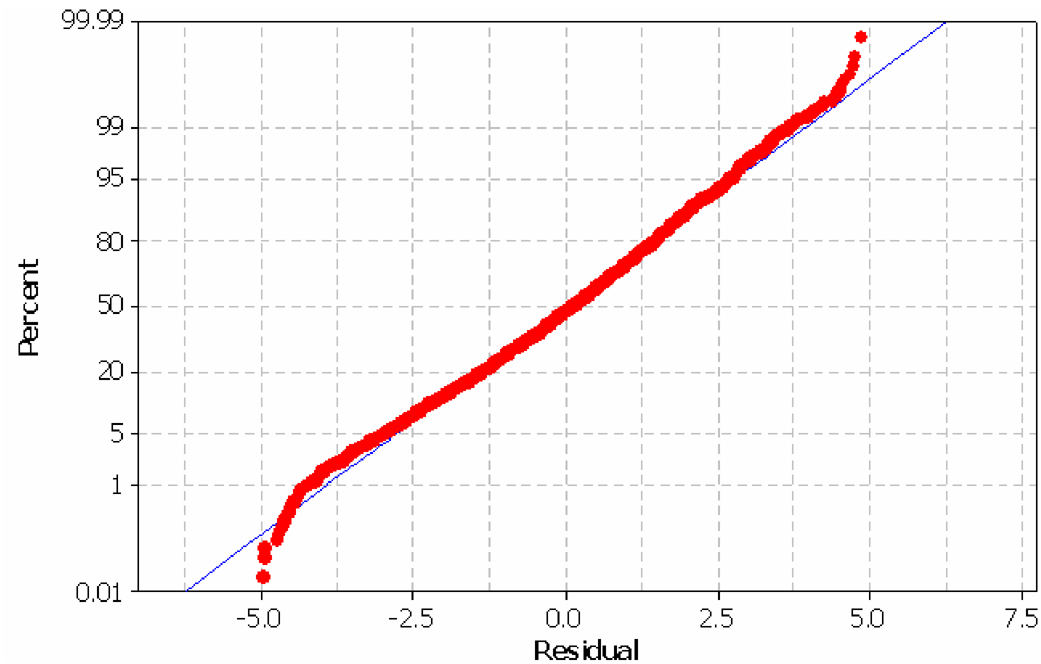

Figure 20. The normal plot of residuals

Table 9. ANOVA results of response time w.r.t all factors and their interactions

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Color combination	3	2304.80	2334.45	778.15	273.55	0.000*
Flashing	2	118.02	133.16	66.58	23.41	0.000*
Color co×Flashing	6	55.41	55.41	9.23	3.25	0.003*
Age	2	1799.19	1819.98	909.99	319.90	0.000*
Gender	1	2.69	2.66	2.66	0.93	0.334
Age×Gender	2	457.71	462.11	231.06	81.23	0.000*
Error	3207	9122.73	9122.73	2.84		
Total	3223	13860.54				

* significance level = 0.05

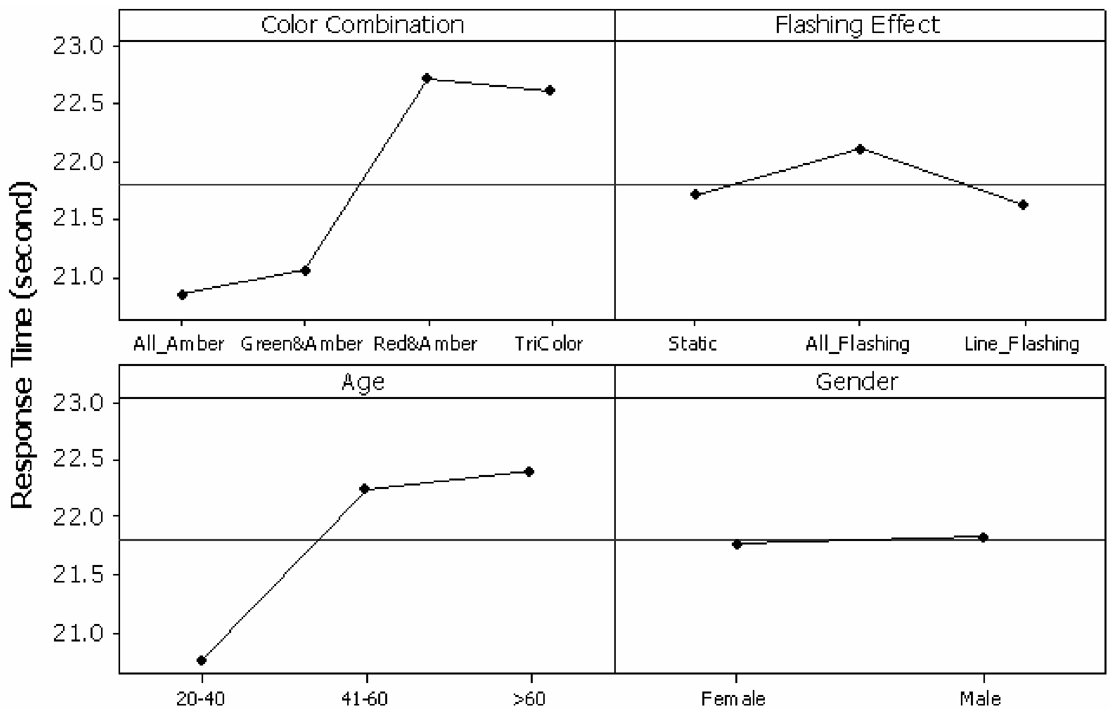

Figure 21. Main effect plots of response time w.r.t. all factors

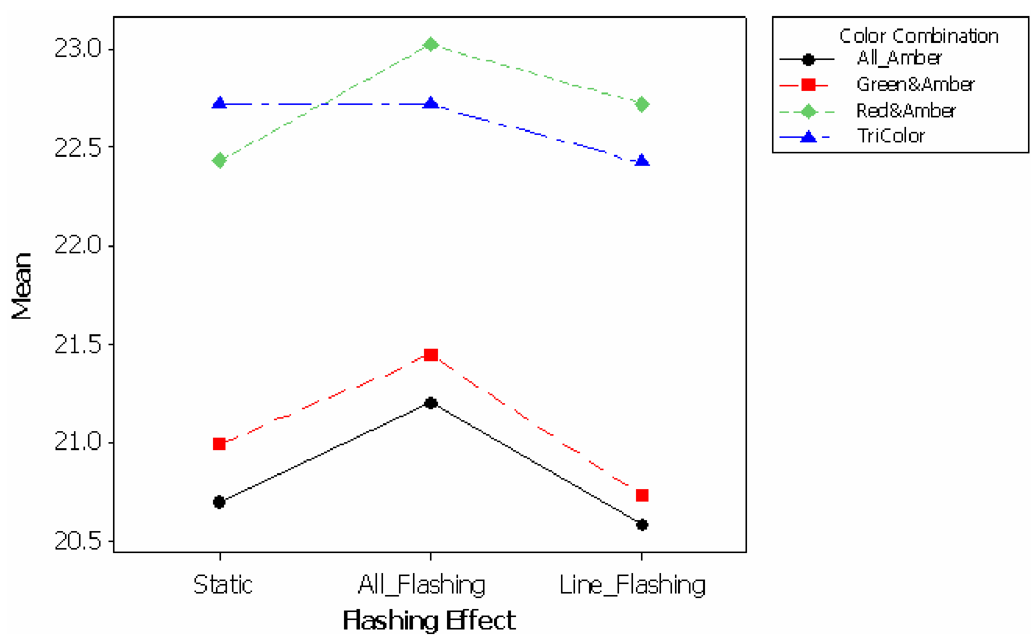

Figure 22. Interaction plots of response time w.r.t. color combination and flashing effect

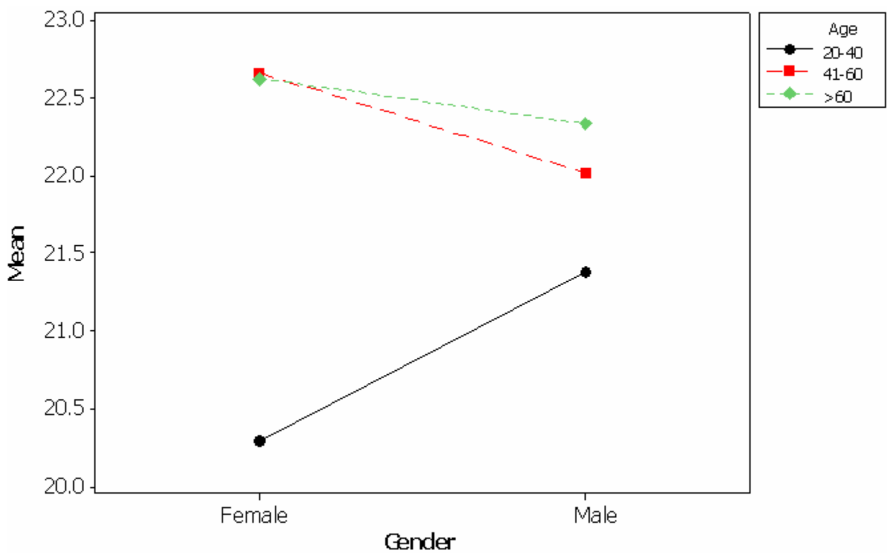

Figure 23. Interaction plots w.r.t. age and gender

Among the four color combinations examined, all amber and green-amber colored messages were not significantly different (p=0.1197) and both resulted in significantly faster response time than others. Subjects took longer to respond to messages containing the red color. This agrees with the survey findings where subjects preferred all amber and green-amber messages over red-amber and tricolor messages.

As to the flashing effect, subjects responded significantly faster to static or one-line flashing messages than those all flashing messages. This also agrees with the survey findings. The difference between static and one-line flashing was not significant (p=0.4092). According to the interaction plots (Figure 22), the fastest-responded VMS messages were those with no flashing or one-line flashing displayed in all-amber color. Younger subjects responded faster with higher accuracy (see Table 10) than older subjects. Female and male subjects did not exhibit a significant difference (p=0.5128) but female subjects responded slightly faster with higher accuracy. In addition, younger female subjects responded faster than males while older male subjects responsed faster than females.

Table 10. Accuracy and response time statistics by age and gender

Age Group	Gender	Accuracy	Response Time (second)	
			Mean	Std. Dev.
20 – 40 yrs.	Female	97%	20.30	0.08949
	Male	95%	21.38	0.09006
41 – 60 yrs.	Female	96%	22.66	0.07773
	Male	95%	22.01	0.07859
> 60 yrs.	Female	92%	22.62	0.07528
	Male	88%	22.33	0.07744

Some past studies indicated that drivers' responses could be affected by message content (*7, 11, 12*). It might be presumed that subjects would respond differently to different message contents tested in the experiments. An additional ANOVA was conducted to investigate this issue by grouping the messages into *warning*, *advisory*, and *fake* messages. The ANOVA results (see Table 11) showed that message content was a significant factor at the 0.05 significance level. Main effect plots are shown in Figure 24. Among messages examined, all subjects responded faster to the warning message which was message 1 in the experiment. The advisory messages (2 and 3) and the fake messages (4) all resulted in slower response times. The results suggested that subjects paid special attention to warning messages and responded faster to them than to other messages. In contrast to the three test messages where only one message content was used in each of them, fake messages were made up of message lines extracted from the three test messages. It is reasonable to presume that it could demand more mental efforts to process and recognize a fake message than regular test messages.

Table 11. ANOVA results of response time for message content

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Message content	2	39.558	39.558	19.779	4.61	0.010*
Error	3221	13820.982	13820.982	4.291		
Total	3223	13860.540				

* significance level = 0.05

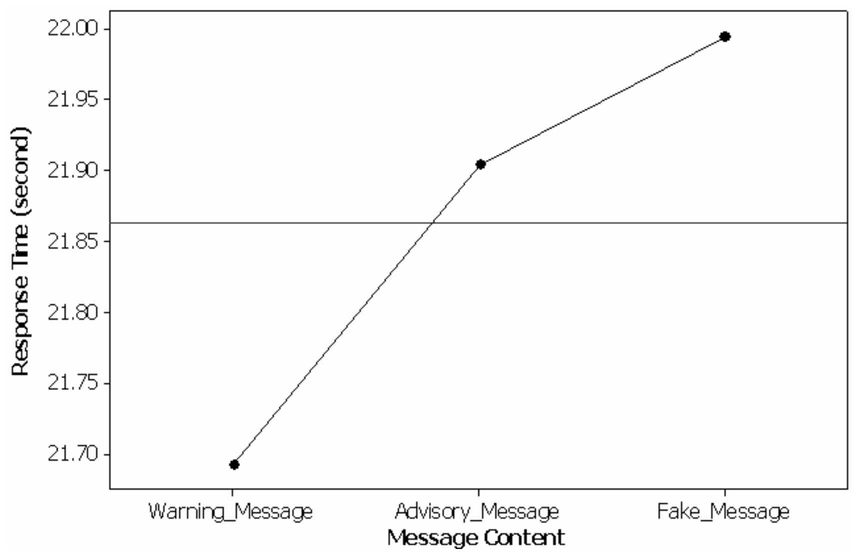

Figure 24. Main effect plots of response time w.r.t. message content

The effects of color combination, flashing, and their interaction were further investigated per each of the warning and advisory messages. Responses to fake messages were not included in the investigation. A set of ANOVA tables were generatd in Tables 12 and 13. Main effect plots and interaction plots are given in Figures 25 through 28. Color combination was the only significant factor for the warning message. All-amber and green-amber colored warning messages resulted in much faster response than others. When looked at the advisory messages, both color combination and flashing effect were significant but not their interaction. All-amber and green-amber colored advisory messages with no flashing or just one-line flashing resulted in much faster response than others. Any messages containing red took longer to respond to. As to the flashing effect, messages with the entire frame flashing resulted in longer response times. Static and one-line flashing messages resulted in faster response regardless of the message type.

Table 12. ANOVA results of response time for warning message

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Color combination	3	921.498	921.134	307.045	97.55	0.000*
Flashing effect	2	15.130	15.081	7.540	2.40	0.092
Color co×Flashing	6	8.326	8.326	1.388	0.44	0.852
Error	821	2584.232	2584.232	3.148		
Total	832	3529.186				

* significance level = 0.05

Table 13. ANOVA results of response time for advisory message

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Color combination	3	1749.84	1750.44	583.48	187.05	0.000*
Flashing effect	2	27.65	27.61	13.81	4.43	0.012*
Color co×Flashing	6	10.29	10.29	1.72	0.55	0.770
Error	1622	5059.53	5059.53	3.12		
Total	1633	6847.32				

* significance level = 0.05

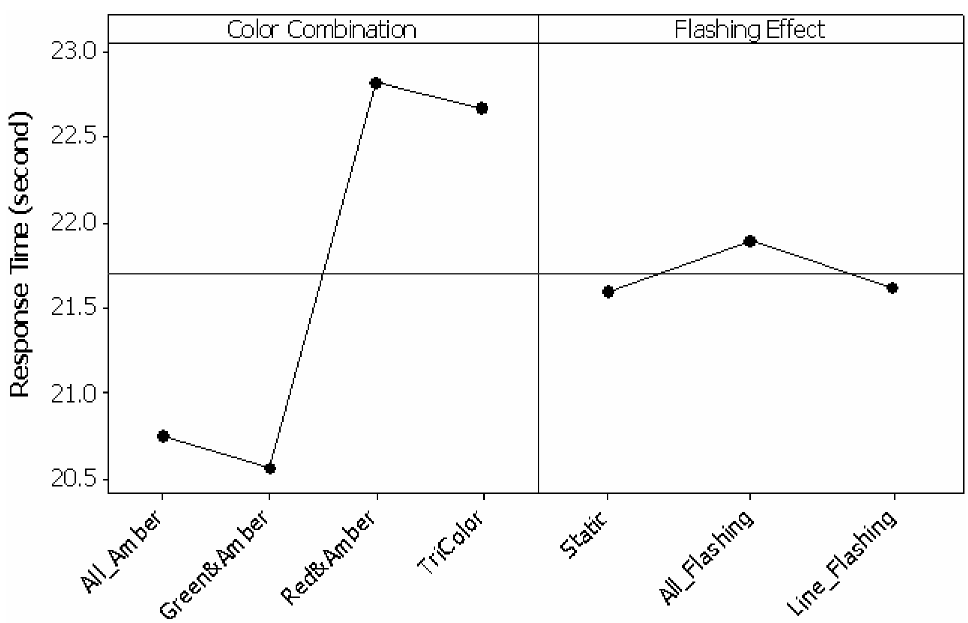

Figure 25. Main effect plots of response time for warning message

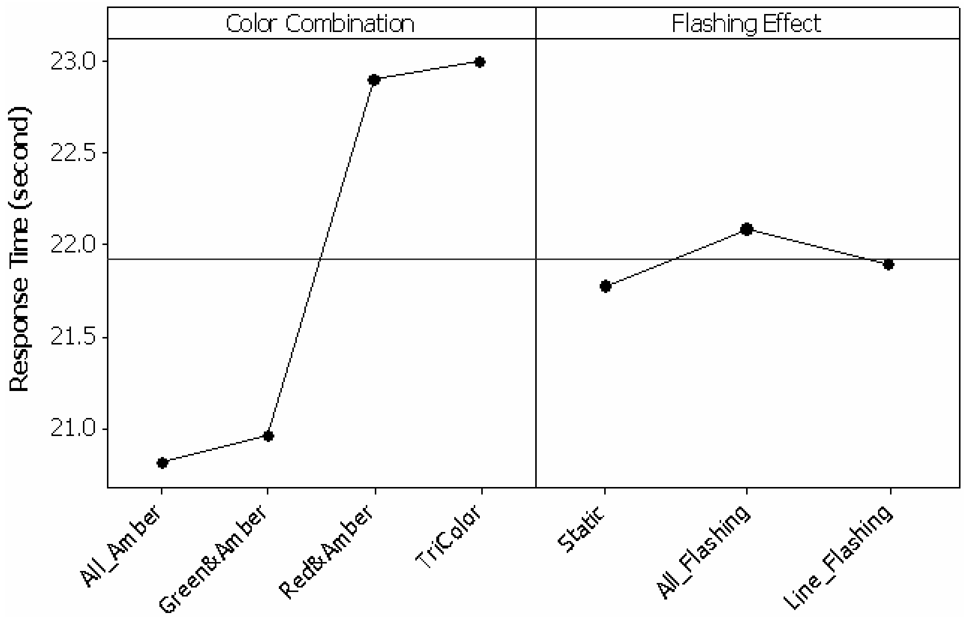

Figure 26. Main effect plots of response time for advisory message

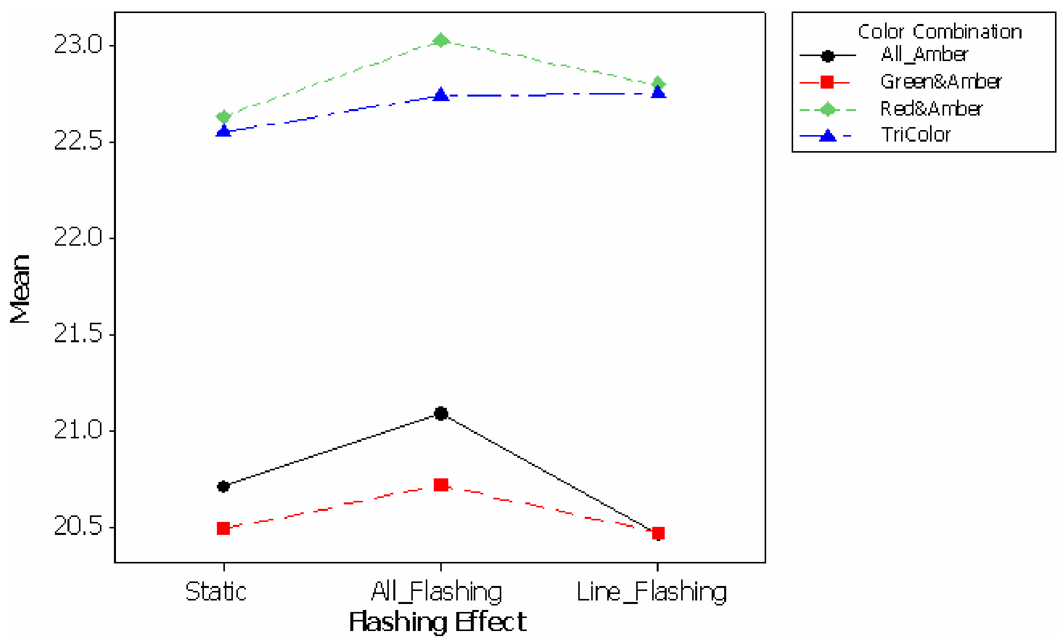

Figure 27. Interaction plots of response time for warning message

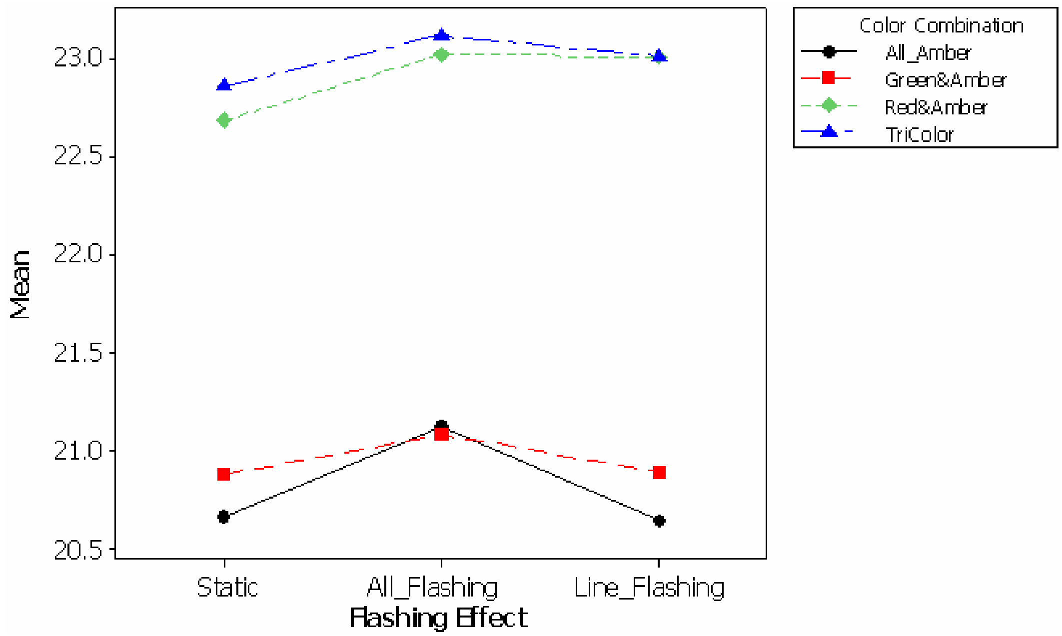

Figure 28. Interaction plots of response time for advisory message

A regression analysis was carried out to investigate the correlation between subjects' response times and accuracies. Regression statistics and ANOVA results are shown in Table 14 while the regression plot is given in Figure 29. No meaningful correlation was found between response time and accuracy (R^2 = 0.2%), i.e., accuracy is independent of response time.

Table 14. Regression statistics and ANOVA results for response time vs. accuracy

Predictor	Coef	SE Coef	T	P
Constant	22.758	3.083	7.38	0.000*
Accuracy	-0.00992	0.03432	-0.29	0.774

S = 1.45788 R-Sq = 0.2% R-Sq(adj) = 0.0%

Source	DF	SS	MS	F	P
Regression	1	0.178	0.178	0.08	0.774
Error	34	72.264	2.125		
Total	35	72.442			

* significance level = 0.05

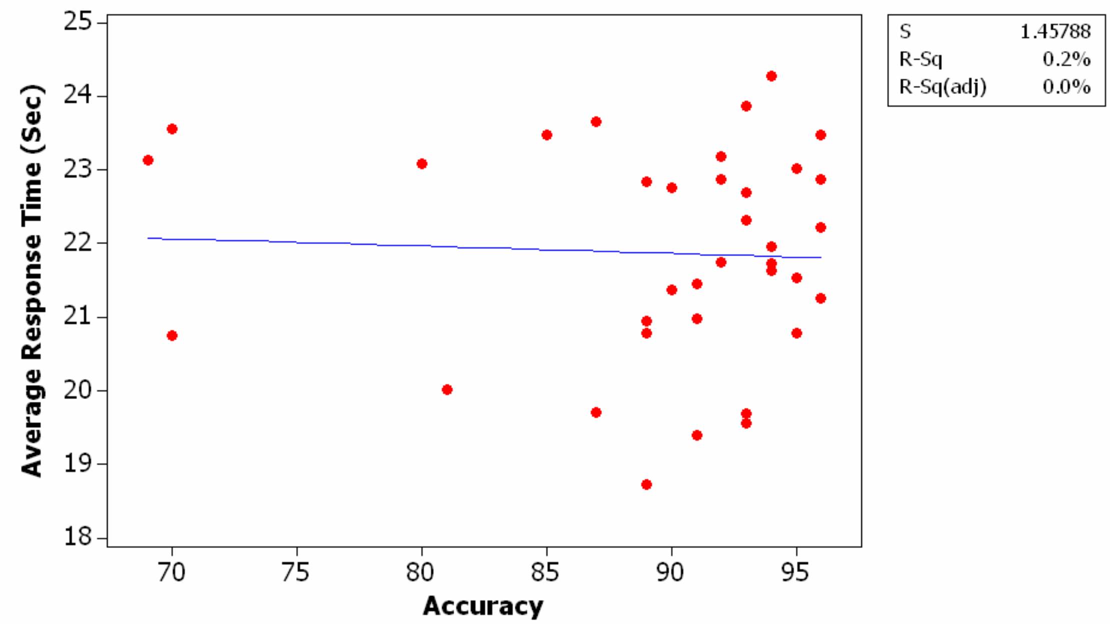

Figure 29. Regression plot for response time vs. accuracy

5.3 Field Study

In the field study, individual driver's response time to the VMS message was collected and calculated via a frame-by-frame audio/video analysis. A direct comparison between individual subject's response times found in the lab experiment and measured in the field study is shown in Figure 30. The average response times measured in real driving were differed approximately by a constant from those obtained in the lab simulation experiment for all test subjects. The average difference was 12.01 seconds ranging 9.20 seconds and 16.53 seconds. It shall be noted that the starting time in response time measurement were different between lab experiment and field study. The starting point of response time in the lab experiment was set as the starting of the simulation video clip while it was set by a reference point along the way to the VMS in the field study. Without this difference, the response times between lab experiments and field studies could be very similar across all subjects.

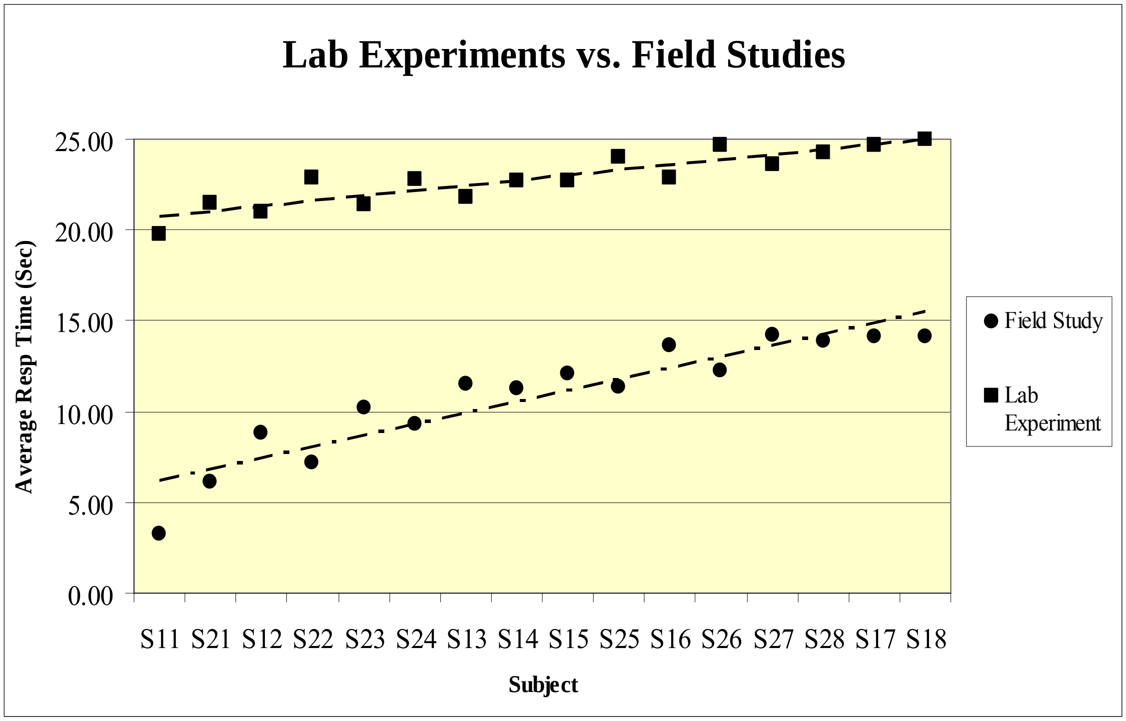

Figure 30. Subject-by-subject comparisons between lab experiments and field studies

To investigate the correlation between lab experiments and field studies, a simple linear regression (SLR) analysis was conducted with the mean response times obtained from the field study as the dependent variables. Regression statistics and ANOVA results are shown in Table 15. The regression equation was found as:

$$\text{Field Study} = -28.1 + 1.70 \text{ Lab Experiment} \tag{5.1}$$

where:

Field Study = a subject's mean response time in the field study (in seconds); and

Lab Experiment = a subject's mean response time in the lab experiment.

With an adjusted R^2 value of 59.7%, the ANOVA results showed that the regression model was significant at the 0.05 significance level. The regression plot is given in Figure 31. The plot showed that all observed data were within 95% prediction interval and indicated a positive, linear relationship between lab experiment and field study. The simple linear regression analysis suggested a moderately strong linear correlation between lab experiment and field study.

Table 15. SLR statistics and ANOVA results for lab experiment vs. field study

Predictor	Coef.	SE Coef.	T	P
Constant	-28.072	8.094	-3.47	0.004*
Lab Experiment	1.7026	0.3534	4.82	0.000*

S = 2.03288 R-Sq = 62.4% R-Sq(adj) = 59.7%

Source	DF	SS	MS	F	P
Regression	1	95.913	95.9135	23.21	0.000*
Error	14	57.856	4.1326		
Total	15	153.770			

* significance level = 0.05

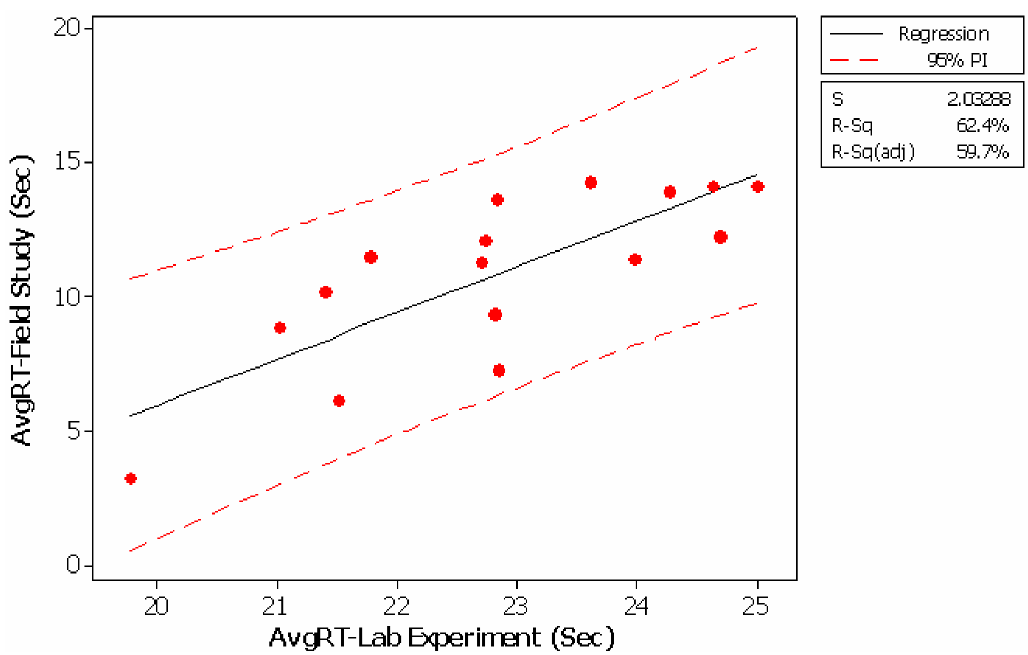

Figure 31. SLR plot for lab experiment vs. field study

To seek a better prediction model, additional factors were considered in the investigation including subjects' actual age, gender, and response accuracy. A multiple linear regression

(MLR) was conducted employing the best subset technique. Table 16 showed that seven feasible models were generated through the best subset procedure. Among the seven models, a model with high adjusted R^2, low s, and a C-p value close to the number of variables in the model would be consider as the best model. Model number 5 fulfilled all three requirements and was chosen as the predicting model. Table 17 exhibits regression statistics and ANOVA results of the selected model. Three predictors, mean response time of lab experiment, subject's actual age, and subject's response accuracy were employed in the model. The predicting model was:

$$\text{Field Study} = -45.1 + 1.60\ \text{Lab Experiment} + 0.0598\ \text{Age} + 17.9\ \text{Accuracy} \tag{5.2}$$

The model had a moderately strong adjusted R^2 value = 75.5% and was significant at the 0.05 significance level, indicating a good predicting power.

Table 16. Best subset regression results

Response is Field Study						Lab Experiment	Age	Gender	Accuracy
no.	Vars	R^2	R^2(adj)	C-p	S				
1	1	62.4	59.7	9.3	2.0329	X			
2	1	45.9	42.0	18.6	2.4386		X		
3	2	74.0	70.0	4.7	1.7528	X			X
4	2	68.7	63.8	7.7	1.9256	X		X	
5	3	80.4	75.5	3.1	1.5860	X	X		X
6	3	75.3	69.1	6.0	1.7800	X		X	X
7	4	80.5	73.5	5.0	1.6492	X	X	X	X

Table 17. MLR statistics and ANOVA results of the best model

Predictor	Coef	SE Coef	T	P
Constant	-45.11	11.29	-4.00	0.002*
Lab Experiment	1.5959	0.3703	4.31	0.001*
Age	0.05980	0.03037	1.97	0.072
Accuracy	17.940	6.471	2.77	0.017*

S = 1.58602 R-Sq = 80.4% R-Sq(adj) = 75.5%

Source	DF	SS	MS	F	P
Regression	3	123.584	41.195	16.38	0.000*
Error	12	30.185	2.515		
Total	15	153.770			

* significance level = 0.05

5.4 Overall Discussion

As described earlier, this study employed three approaches, questionnaire surveys, lab experiments, and field studies. Findings found from one approach provided support and validation to findings found from others. Results obtained from the lab experiment supported the preference statistics found from the survey. In particular, the all amber and green-amber colored messages preferred by most subjects in the survey had the shortest response times in the lab experiment; while the least preferred red-amber message had the longest response times (see Table 18.)

Table 18. Overall response statistics from survey and lab experiment w.r.t. color combination

Color Combination	Survey (Preference %)	Experiment (Mean RT in Seconds)
All Amber	33.55%	20.83
Green & Amber	38.16%	21.05
Red & Amber	7.24%	22.73
Tricolor	21.05%	22.62

As for the flashing effect (Table 19), a static or one line flashing message was the most preferred in the survey and resulted in shorter response times in the lab experiment while the least preferred whole frame flashing message had longer response time.

Table 19. Overall response statistics from survey and lab experiment w.r.t. flashing effect

Flashing	Survey (Preference Percentage)	Experiment (Mean RT in Seconds)
Static	54.29%	21.71
One Line Flashing	36.50%	21.62
All Flashing	9.21%	22.10

When subjects' demographics were examined, the most preferred color combination by all age groups and both genders had the shortest response times (see Table 20.) Same results were also found regarding the flashing effect (see Table 21.) Overall, it should be noted that the most preferred choice in the survey resulted in the shortest response times in the lab experiment. It suggests that subjects' preferences in the survey are correlated with their responses in the experiment and findings from the questionnaire survey are supported by those from lab experiment.

Table 20. Demographic-specific response statistics from survey and lab experiment w.r.t. color combination

Color Combination		Survey (Preference Percentage)				Experiment (Mean RT in Seconds)			
		All Amber	Green & Amber	Red & Amber	Tricolor	All Amber	Green & Amber	Red & Amber	Tricolor
Age Group	20-40	27.08%	37.50%	6.25%	29.17%	19.58	19.70	22.09	22.00
	41-60	47.92%	39.58%	8.33%	4.17%	21.33	21.32	23.20	23.36
	61-above	27.09%	39.58%	6.25%	27.08%	21.47	21.43	23.37	23.36
Gender	Male	37.50%	33.33%	5.56%	23.61%	20.84	20.82	22.81	22.89
	Female	30.56%	44.44%	8.33%	16.67%	20.73	20.80	22.94	22.89

Table 21. Demographic-specific response statistics from survey and lab experiment w.r.t. flashing effect

		Survey (Preference Percentage)			Experiment (Mean RT in Seconds)		
		Static	One Line Flashing	All Flashing	Static	One Line Flashing	All Flashing
Age Group	20-40	41.67%	45.83%	12.50%	20.73	20.80	20.99
	41-60	66.67%	29.17%	4.17%	22.22	22.26	22.43
	61-above	66.67%	27.08%	6.25%	22.22	22.35	22.64
Gender	Male	52.78%	33.33%	13.89%	21.69	21.84	21.99
	Female	61.11%	34.72%	4.17%	21.72	21.76	22.04

When comparing the response times between lab experiment and field study, three models were developed. They were subject-by-subject model, simple linear regression (SLR) model, and multiple linear regression (MLR) model. The first model directly compared the mean response times between lab experiment and field study on a subject-by-subject basis. It found that the difference between these two approaches was nearly a constant. The second model was to correlate response time to VMS in real driving and in lab driving simulation. It found that there was a positive, linear, significant correlation between them with a moderately strong adjusted R^2. The last model considered several predictors in order to develop a better model. The model was found with a good adjusted R^2 value employing the mean response time of lab experiment, subject's actual age, and subject's response accuracy as predictors. It suggested that response obtained from simulated driving lab experiment could be used to predict drivers' response in actual driving.

6. CONCLUSIONS

A comprehensive, human factors study was carried out to assess drivers' responses and preferences to various combinations of VMS messaging features. It incorporated three methodological approaches: a questionnaire survey, a lab experiment, and a field study. The study found new insights and important implications regarding the design and display of VMS messages that would not have been available through the use of a single approach.

Questionnaire surveys suggested a VMS message to be a one-frame message with minimum flashing, very specific wording, no abbreviation, and displayed in solid amber or green-amber color combination. Lab simulation experiments found that a static or one-line flashing message displayed in solid amber or green-amber color combination demanded less response time. Experiment results from the lab experiment supported the findings obtained from the survey. In particular, the all amber and green-amber colored messages preferred by most in the survey had the shortest response time in the lab experiment while the least preferred red-amber message had the longest response time. Also, a static or one-line flashing message preferred by the majority in the survey resulted in shorter response time in the lab experiment while the least preferred whole frame flashing message had the longest response time. In general, it found that those message preferred by the majority usually resulted in shorter response times in the lab experiment.

Results from field studies were correlated with those obtained from laboratory experiments. Through a direct one-on-one comparison, it found that the mean response times to the same VMS message in real driving and in lab setting differed approximately by a constant across all test subjects. If the difference in the starting time set up was excluded, the two response times could be very close across all subjects. Through a simple linear regression analysis, a moderately strong linear correlation was found between lab experiment and field study. Based on this correlation, additional factors were considered in a multiple linear regression approach to seek a stronger correlation model. Using mean response time of lab experiment, subject's actual age, and subject's response accuracy as predictors, a stronger model with good predicting power was found. The above-mentioned findings provided evidence to support the presumptions that the video-based driving simulation employed in the laboratory experiment is an effective and economical means to gauge drivers' responses to and comprehension of VMS messaging in real driving.

Among the three approaches, gender effects were nearly negligible while the age effect was more noticeable. Younger subjects' performances are somewhat different than those of older subjects. It found that younger subjects took less time to respond to the VMS messages with higher accuracy than older subjects. The age effect findings agree with the results from past studies (*2, 3, 16, 17, 35*). All subjects responded faster to messages with less flashing displayed in either solid amber or green-amber colors. The findings indicate that the age differences did exist and further studies are needed to address the needs of elder drivers.

Overall, the present findings suggest some practical implications with a specific set of VMS message display features that were preferred by drivers. These preferred VMS features could help enhance drivers' comprehension of and response to VMS messages. Although this study was focused only on those Daktronics' Vanguard® VMS systems currently in service in Rhode Island, it could be extended to similar systems in other states since the LED-based VMS adopted nearly identical technology. According to the latest VMS installation data provided by Daktronics, the VMS systems studied here are in fact employed by 23 states. The research found here could benefit highway authorities in these states. These findings could serve as a foundation or baseline for message display and design improvement on most VMSs. They could help traffic engineers and highway management design driver-friendly VMS signs that could be noticed, understood and responded to in a more timely fashion. It could eliminate confusion and frustration for the drivers and result in a safer and pleasant daily driving experience.

When considering theoretical implications of this study, the findings indicated that the demographic differences need to be examined in further studies, especially for the elder and ESL populations. This study could also provide a baseline for future works that explore other advanced features of VMS messaging. These features could include but are not limited to graphics, pictograms, and animations. In summary, the methodological approaches employed in this study were based on a sound three-way scientific approach that could be replicated by other. The research was cost-effective and reliable in its designs and execution. Studies based on this research could yield further cross-validated conclusions and many beneficial applications.

7. REFERENCES

1. Daktronics. *Vanguard® VMS ITS Designer's Tool Kit (Rev 02)*, Daktronics, Brookings, S.D., 2005.

2. Wang, J.H., and Y. Cao. A Human Factors Study on Message Design of Variable Message Signs. *International Journal of Industrial Engineering*, Vol. 10, No. 4, 2003, pp. 339-344.

3. Wang, J.H., and Y. Cao. Assessing the Design and Display of Messages on Portable Variable Message Signs, *Proceedings of the 84th Annual Meeting of Transportation Research Board*. CD-ROM, Paper No. 05-2186, 2005, pp. 1-23.

4. US Census 2000. http://www.census.gov/population/www/cen2000/briefs.html. Accessed July 12, 2005.

5. ITS Deployment Tracking: 2002 Survey Results. http://itsdeployment2.ed.ornl.gov/its2002/default.asp. Accessed July 15, 2005.

6. Benson, B.G. Motorist Attitudes about Content of Variable-Message Signs, In *Transportation Research Record: Journal of the Transportation Research Board, No. 1550*, TRB, National Research Council, Washington, D.C., 1996, pp. 48-57.

7. Peng, Z.-R. et al. Motorist Response to Arterial Variable Message Signs. In *Transportation Research Record: Journal of the Transportation Research Board, No.1899*, TRB, National Research Council, Washington, D.C., 2004, pp. 55-63.

8. Emmerink, R. et al. Variable Message Signs and Radio Traffic Information: An Integrated Empirical Analysis of Drivers' Route Choice Behaviour. *Transportation Research – A*, Vol. 30, No. 2, 1996, pp. 135-153.

9. Yim, Y., and J.L. Ygnace. Link Flow Evaluation Using Loop Detector Data: Traveler Response to Variable Message Signs. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1550*, TRB, National Research Council, Washington, D.C., 1996, pp. 58-64.

10. Chatterjee, K. et al. Driver Response to Variable Message Sign Information in London. *Transportation Research – C*, Vol. 10, No. 2, 2002, pp. 149-169.

11. Wardman, M. et al. Motorist response to variable message signs: a stated preference investigation. *Transportation Research – C*, Vol. 5, No. 6, 1997, pp. 389-405.

12. Peeta, S. et al. Content of Variable Message Signs and On-line Driver Behavior. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1725*, TRB, National Research Council, Washington, D.C., 2000, pp. 102-108.

13. Bonsall, P.W., and I. Palmer. The Impact on Route Choice of Direction Advice and Guidance from DMS Signs. *Colloquium on Dynamic Control of Strategic Inter-urban Road Networks*, London, England, 1995, pp. 1-10.

14. Lai, K.H., and W.G. Wong. The Influence of Traffic Information Presentation Format on Driver Behaviour. *Proceedings of the Third Hong Kong Society for Transportation Studies*, 1998, pp. 353-358.

15. Boyle, L.N., and F. Mannering. Impact of Traveler Advisory Systems on Driving Speed: Some New Evidence. *Transportation Research – C*, Vol. 12, 2004, pp. 57-72.

16. Guerrier, J.H., and J.A. Wachtel. A Simulator Study of Driver Response to Changeable Message Signs of Differing Message Length and Format. *Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Aspen, Colorado, 2001, pp. 164-165.

17. Guerrier, J.H. et al. Drivers' Response to Changeable Message Signs of Differing Message Length and Traffic Conditions. *Contemporary Ergonomics*, P.T. McCabe (Ed.), London: Taylor and Francis, 2002, pp. 223-228.

18. Metaxatos, P., and S. Soot. Evaluation of motorist's ability to recall the message content of portable changeable message signs in highway work zones. *Journal of Transportation Research Forum*, Vol. 40, No. 1, 2001, pp. 129-141.

19. Nsour, S.A. IVHS and the Elderly Driving. *Traffic Congestion and Traffic Safety in the 21st Century: Challenges, Innovations, and Opportunities*, Chicago, Illinois, 1997, pp. 333-339.

20. Harder, K.A. et al. *The Effectiveness and Safety of Traffic and Non-traffic Related Messages Presented on Changeable Message Signs (CMS)*. Publication MN/RC-2004-27, Minnesota Department of Transportation, 2003.

21. Ulfarsson, G.F., V.N. Shankar, and P. Vu. The Effect of Variable Message Signs on the Relationship Between Mean Speeds and Speed Deviations. *Proceedings of the 81st Annual Meeting of Transportation Research Board*. CD-ROM, 2002, pp. 29.

22. Armstrong, J. D., and J.E. Upchurch. Human factors design considerations for variable message freeway signs, *Journal of Transportation Engineering*, Vol. 120, No.2, 1994, pp. 264-282.

23. Durkop, B.R., and C.L. Dudek. Texas Driver Understanding of Abbreviations for Changeable Message Signs. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1748*, TRB, National Research Council, Washington, D.C., 2001, pp. 87-95.

24. Hustad, M.W., and C.L. Dudek. Driver Understanding of Abbreviations on Changeable Message Signs in New Jersey. In *Transportation Research Record: Journal of the Transportation Research Board, No.1689*, TRB, National Research Council, Washington, D.C., 1999, pp. 46-52.

25. Dudek, C.L. and G.L. Ullman. *Variable Message Sign Operations Manual*. Publication FHWA-NJ-2001-10, Texas Transportation Institute, 2001.

26. Dudek, C.L. and G.L. Ullman. Flashing Messages, Flashing Lines, and Alternating One Line on Changeable Message Signs. In *Transportation Research Record: Journal of the*

Transportation Research Board, No. 1803, TRB, National Research Council, Washington, D.C., 2002, pp. 94-101.

27. Dudek, C.L. et al. Dynamic Visual Displays. *Human Factors Design of Dynamic Visual and Auditory Displays for Metropolitan Traffic Management*, Vol. 2, Publication FHWA-RD-81/040. Texas Transportation Institute, 1981.

28. Staplin, L. et al. *Highway Design Handbook for Older Drivers and Pedestrians*. Publication FHWA-RD-01-103. FHWA, U.S. Department of Transportation, 2001.

29. Dudek, C.L. and R.D. Huchingson. *Manual on Real-Time Motorist Information Displays*. Publication FHWA-IP-86-16. FHWA, U.S. Department of Transportation, 1986.

30. Federal Highway Administration (FHWA). *MUTCD 2003-Manual on Uniform Traffic Control Devices*, U.S. Department of Transportation, Washington, D.C., 2003.

31. Kemeny, A., and F. Panerai. Evaluating Perception in Driving Simulation Experiments, *Trends in Cognitive Sciences*, Vol. 7, No. 1, 2003, pp. 31-37.

32. Shinar, D. *Psychology on the Road: the Human Factors in Traffic Safety*, New York: John Wiley & Sons, 1978.

33. Abdel-Aty, M.A. et al. *Impact of ATIS on Driver's Travel Decisions: a Literature Review*, Publication UCD-ITS_RR-92-7, UC Davis, Institute of Transportation Studies, Davis, CA, 1992.

34. Bonsall, P.W. et al. Validating the Results of a Route Choice Simulator, *Transportation Research – C*, Vol. 5, No. 6, 1997, pp. 371-387.

35. Lee, H.C. et al. Assessing the Driving Performance of Older Adult Drivers: On-road versus Simulated Driving, *Accident Analysis and Prevention*, Vol. 35, 2003, pp. 797-803.

36. Waller, P.F. The Older Driver, *Human Factors*, Vol. 33, No. 5, 1991, pp. 499-505.

37. Carr, D. The Older Adult Driver. *American Family Physician*, Vol. 61, Issue 1, 2000, p. 141.

38. Chrysler, S. et al. A Comparison of Traffic Sign Comprehension Using Static, Dynamic and Interactive Presentation Media, *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2003, pp. 219-220.

39. Edwards, C.J. et al. Older and Younger Driver Performance at Complex Intersections: Implications for Using Perception-Response time and Driving Simulation, *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2003, pp. 33-38.

APPENDIX A: EXPERIMENT DATA

Table A1. Experiment data example

Subject	Age	Gender	Message Content	Color Comb.	Flashing Effect	Response Time (second)	Response Key	Accuracy
Greene	1	1	1	1	0	20.71875	49	0
Field	3	2	4	4	2	22.29688	52	1
:	:	:	:	:	:	:	:	:

Note: 1. In column "Age", 1 stands for 20-40 year old, 2 stands for 40-60 year old, and 3 stands for over 60 year old.

2. In column "Gender", 1 stands for female and 2 stands for male.

3. In column "Message Content", 1 stands for message 1, 2 stands for message 2, 3 stands for message 3, and 4 stands for fake messages.

4. In column "Color Comb.", 1 stands for amber, 2 stands for green & amber, 3 stands for red & amber, and 4 stands for tricolor.

5. In column "Flashing Effect", 0 stands for static, 1 stands for all flashing, and 2 stands for line flashing.

6. In column "Response Key", 49 stands for button "1", 50 stands for button "2", 51 stands for button "3", and 52 stands for button "4".

7. In column "Accuracy", 0 stands for the incorrect response to a VMS message and 1 stands for the correct response to a VMS message.