A DESIGN OF EXPERIMENT APPROACH TO STUDY THE DISPLAY OF VARIABLE MESSAGE SIGNS

FHWA-RIDOT-RTD-02-7

DECEMBER 2002

Jyh-Hone Wang, Ph.D.

Christopher Hunter, Ph.D.

Yong Cao, M.S.

Department of Industrial and Manufacturing Engineering

University of Rhode Island

Kingston, RI 02881

Sponsored By

Rhode Island Department of Transportation

Research and Technology Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-RIDOT-RTD-02-7	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle A DESIGN OF EXPERIMENT APPROACH TO STUDY THE DISPLAY OF VARIABLE MESSAGE SIGNS		**5. Report Date** December 2002 **6. Performing Organization Code**
7. Author(s) Jyh-Hone Wang, Ph.D. Christopher Hunter, Ph.D. Yong Cao, M.S.		**8. Performing Organization Report No.**
9. Performing Organization Name and Address University of Rhode Island Department of Industrial and Manufacturing Engineering Kingston, RI 02881		**10. Work Unit No.** **11. Contract or Grant No.** SPR-226-2258
12. Sponsoring Agency Name and Address Rhode Island Department of Transportation Two Capital Hill Providence RI 02903-1124		**13. Type of Report and Period Covered** Final Report, May 2001 – October 2002 **14. Sponsoring Agency Code**
15. Supplementary Notes Colin A. Franco, P.E., RIDOT Research and Technology Development, Managing Engineer		

16. Abstract

This report describes a study on the design and display factors of variable message signs (VMSs). The two-phase study was conduct through a series of blocked-factorial experiments. Computer generated VMS images, merged with a driver's view driving video, were projected onto a screen in front of the test vehicle. Subjects in the driver's seat were required to make proper responses signaling their comprehension of the VMS stimuli. Eighteen subjects balanced by age and gender participated the experiments. Phase I investigated the effects of discretely displayed VMSs' font size, font color, subjects' age, gender, and their interactions. It found that font color, drivers' age, and gender significantly affected response time. Green and 5 × 7 dot matrix font were the best font color and font size respectively. Older drivers responded the fastest among the three age groups but with the lowest accuracy. No significant correlations were found between subjects' response time and accuracy. Response times of different subjects were significantly different, but the effects of font color and size were consistent among different subjects. Adopting specific recommendations based on the phase I findings, phase II studied the influences of display format, number of message lines, weather, driving lane, and their interactions. It found that discretely displayed messages took less response time than sequentially displayed ones. Single-line messages were better than multiple-line ones. Motorists could better view VMSs in sunny days, and better view VMSs when driving in the outer lane. Older drivers exhibited slower response and less accuracy than younger drivers; females exhibited slower response but higher accuracy than males.

17. Key Words	**18. Distribution Statement**
Variable message signs, human factors, design of experiment.	No restriction. Available from: Rhode Island Department of Transportation The Research and Technology Development Two Capitol Hill, Room 013 Providence RI 02903-1124 University of Rhode Island Department of Industrial and Manufacturing Engineering Kingston, RI 02881

19. Security Classification (For this report) None	**20. Security Classification (For this page)** None	**21. No. of Pages** 52	**22. Price**

ACKNOWLEDGEMENTS

This research was made possible with the funding and assistance from the Rhode Island Department of Transportation.

Special thanks to all those at the Rhode Island Department of Transportation listed below.

Name	*Organization*
Colin A. Franco, P.E.	Research and Technology Development, Managing Engineer
Francis Manning, P.E.	Research and Technology Development
Joseph Schall Jr.	Transportation Management Center
Robert Rocchio	Traffic and Safety Management Section
Michael Sock	Research and Technology Development

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

This report describes the research on the design and display of variable message signs (VMSs). VMSs, also called changeable message signs (CMSs) or dynamic message signs (DMSs), are programmable traffic control devices that display messages composed of letters, symbols, or both. The information displayed on a VMS can come from a variety of traffic monitoring and surveillance systems. They can be changed by a system monitor through remote control or automatic controls that can "sense" the conditions requiring special messages. VMSs are usually applied to deal with the following five categories of operational problems found on highways: 1) recurring problems, 2) nonrecurring problems, 3) environmental problems, 4) special event traffic problems, and 5) special operational problems (*1*). With more sophisticated technologies, VMSs are gaining widespread uses in intelligent transportation systems (ITS), especially in informing motorists of various situations particularly along more congested traffic corridors. The success of ITS is ultimately contingent on the development and delivery of VMS messages with good visibility, legibility and understandability. This is critical especially in high-volume, high risk, and construction/repair zones. In the State of Rhode Island, eight VMSs are in operation when this study started, with thirteen more scheduled for deployment soon, there is a need to evaluate the design and display of VMSs and their impacts on drivers' responses.

Aiming at improving the design and display of VMSs, a two-phase study sponsored by the Rhode Island Department of Transportation was conducted. Chapter 2 explains the objective and scope of the research. Chapter 3 gives a literature review on past and current researches on VMSs. Chapter 4 outlines the methodology employed in this study. Chapter 5 describes the phase I experiments. Chapter 6 describes the phase II experiments. Chapter 7 lists the recommendations, followed by References and Appendices.

2. OBJECTIVE AND SCOPE OF THE RESEARCH

The purpose of the research is to identify factors in VMS design and display that could impact a driver's reception of displayed messages. It also investigated and assessed the impacts of age and gender on a driver's response to a combination of these factors. With the intention to optimize the identified factors in VMS design and display, a number of human factor experiments were conducted. Employing a blocked factorial experimental design, these experiments studied font size, font color, VMS display format, weather, driving lane, subjects' age, and subjects' gender, each with several settings, and their interactions. The overall objective is seeking an effective VMS display that bridge highway management and drivers to warrant a safer and more efficient driving on Rhode Island highways.

3. LITERATURE REVIEW

A study conducted in northern Virginia assessed motorists' attitudes toward VMSs and the effect of demographic characteristics on these attitudes. In the responses to the survey question regarding how often VMSs influenced their driving, half the respondents replied "often", 40% answered "occasionally", and the others indicated "not at all". In other words, half of the respondents depended regularly on VMSs, which was consistent with attitudes expressed in the focus groups. A survey of more than 500 motorists in the Washington D.C. area obtained the similar results (*2*). It also found that demographic variables such as age, education, income, and gender, appeared to have very little influence on motorists' attitudes toward VMSs. The interview surveys conducted in Paris revealed that 97% of drivers were aware of the existence of VMSs, 62% completely understood the information presented on VMSs, 84% considered the information presented to be useful, and 46% had on at least one occasion diverted in response to the travel time information (*3*).

Many investigations were performed on evaluating the effects of VMSs on drivers' response and behavior. A survey conducted in Finland revealed that 91% of the drivers recalled the posted speed limits, 66% recalled the slippery road sign, and 34% recalled the temperature display, indicating that drivers could recall VMSs better than regular fixed signs (*4*). The research on travelers' response to VMSs in Paris concluded that VMSs alone could affect vehicle diversion significantly (*5*). VMSs were most effective when displayed during periods of increasing congestion, and the responses to VMSs were more significant during morning peak hours than evening peak hours; the longer the queue length posted in VMSs, the more vehicles diverted. Wardman et al assessed the effect of VMSs information on drivers' route choice behavior (*6*). They found that the impact of VMS information depended on: 1) the message content, such as the cause of delay and its extent, 2) local circumstances, such as relative journey times in normal conditions, and 3) drivers' characteristics, such as age, gender, and previous network knowledge. The study investigating the impacts of VMSs in London revealed similar results that incident location and message content were important factors influencing the probability of diversion (*7*). Emmerink et al analyzed the impacts of both radio traffic information and VMSs information on drivers' route choice behavior (*8*). The result showed that women were less likely to be influenced by traffic information, and the level of satisfaction with alternative routes was strongly related to the type and distance of the alternative road. The impacts of these two sources of information on route choice behavior were very similar, and there was a positive correlation between their uses. The investigations on the effects of VMSs for slippery road conditions on drivers' behavior found that, VMSs reduced the mean speed in addition to the decrease caused by adverse road conditions (*9*, *10*). It also found that VMSs had other effects on drivers, such as refocusing of attention to seek cues on potential hazards, testing of the road slipperiness, and generating more careful passing behavior. Using a driving simulator, Comte and Jamson studied drivers' behavior under speed-reducing measures for curves (*11*). Roadside VMSs displaying the advisory speed were found to be effective, particularly in help lowering speeds early in the approach to curves. Garber and Srinivasan's research indicated that VMSs were more effective than the regular speed control signs in reducing speeds in work zones. There were no distinctive differences among different types of vehicles with regard to speed reduction, and the installation of a second VMS was needed in very long work zones (*12*). A similar study at South Dakota concluded that the mean traffic speed in work zones was 4-5 mph lower after the installation of speed monitoring VMSs (*13*). Kraan et al assessed network wide effects of VMSs in Netherlands (*14*). The study showed that the use of VMSs had a positive impact on network performance, and VMSs greatly improved the reliability of the system. Both drivers and operators thought VMSs were useful and the information provided had a significant effect on route choice.

Several researches investigated drivers' cognition of VMSs. The study conducted by Metaxatos involved the examination of factors that affected drivers' ability to recall messages provided via VMSs in highway work zones (*15*). Through a chi-square analysis, time of the day, drivers' age, type of vehicles, and familiarity with the site were identified as relevant factors, and drivers were more likely to recall messages that specify an action rather than the problem itself. Dudek et al evaluated drivers' understanding of abbreviations for VMSs in Texas and New Jersey (*16*, *17*, *18*). Specific recommendations regarding VMS abbreviations were given.

As to the visibility of VMSs, one study indicated that it mainly depends on two factors, the visual capability of drivers and the photometry[1] (*19*). A sign's foreground and background color and its brightness are significant to its overall visibility. The effectiveness of a VMS depends on how much time a driver could spend to read it. This time is determined by the approaching speed, the distance from the sign when it is first noticed, and the legibility distance of the sign. Garvey and Mace investigated the VMS visibility (*20*). Their research intended to optimize VMS components, including the character variables (font, width-to-height ratio, color, and contrast orientation) and the message variables (inter-letter, inter-word, and inter-line spacing). Through a field survey of in-use VMSs and a series of static and dynamic field studies, they recommended the optimum character/message variables for VMS visibility.

Regarding the design and display of VMSs, Armstrong et al conducted a field study in Phoenix, Arizona on human factors design considerations such as legibility distance, target value, and viewing comfort for fiber-optic and LED VMSs on freeways (*21*). It concluded that the optimum number of words contained in a message varied with the VMS technology, the

---

[1] Photometry is the measurement of light, which is defined as electromagnetic radiation which is detectable by the human eye. It is thus restricted to the wavelength range from about 360 to 830 nanometers.

lighting conditions, and the prevailing traffic speed. For VMSs using other technologies or with different character fonts or dimensions, a legibility analysis prior to its implementation should be performed. Dudek's research indicated that on two-sequence messages, up to four words or two units of information per sequence could be displayed at the rate of up to 0.5 second/word without loss of recall. Drivers could see the message cycle twice. Messages longer than four words or two units should be cycled at a speed of at least 1 second/word (*1*). Miller et al strongly recommended that a VMS use no more than two message screen, and a single message screen was preferred (*22*). Dudek and Ullman's study on the dynamic characteristics of VMSs suggested that one frame VMS messages and a line on a one-frame VMS message should not be flashed, and a line on a two-frame VMS message should not be alternated while keeping the other lines the same (*23*). A field study aiming at comparing the visual demand of bilingual VMSs displaying alternating text messages suggested that, the sign displaying alternating bilingual messages was no more demanding than the VMS displaying the same messages simultaneously (*24*).

4. EXPERIMENT APPROACH

This study is consisted of two phases. Phase I examined several factors and their impacts on human/ITS interface using various discretely displayed VMSs. The main factors include font color (per the Specifications for Standard Highway Sign Colors) and font size, and the blocking factors[2] are drivers' age and gender. Based on the findings from phase I, phase II study focused on evaluating the display format of VMSs as viewed by motorists driving under different weather and lane conditions. Main factors include VMS display format, number of message lines, weather, and driving lane. Blocking factors are drivers' age and gender. Both phases employed blocked factorial experimental design to investigate the effects of main factors, blocking factors, and their interactions. The experiments were conducted in a driving simulation lab where an environment of real video scenes was integrated with computer-generated VMSs to simulate a virtual driving experience. Various VMS stimuli were introduced in a random but controlled manner. They were similar to those currently used in Rhode Island.

Figure 1 depicts the experimental setup. A sequence of simulated VMSs was generated using MS Visual Basic, merged with a driving video simulating a driver's view driving at 60 mph. A Sony TRV 330 digital camcorder mounted inside a 2001 Chrysler Voyager was used to take videos in segments of I-90 and I-495 westbound in Massachusetts at about 10 am in April and October 2002. The driving video was input into the computer through an "All-in-Wonder 128 Pro" video card. It was used as the background on the computer monitor. The computer generated VMS stimulus was sent to the computer monitor and thus displayed on the driving video background. They were then output to a computer projector and were projected onto a screen in front of the test vehicle.

[2] A blocking factor is the grouping of subjects according to some pre-existing subject similarity.

A test subject sat in the driver’s seat of a 1998 Ford Taurus sedan located 10 feet from the screen (see Figure 1). The subject saw the displayed VMSs appeared on the screen and gradually increased in size. She/he was required to press the pre-labeled key in a keyboard to signify her/his understanding of a specific message. During the experiment, all lights inside the lab were turned off with only the interior light in the vehicle left on. Subjects’ response time and accuracy were recorded in the MS Access database (see Table A1 in Appendix A), which linked with the MS Visual Basic simulation. The response time to a VMS stimulus was calculated as the difference between the time of a subject’s response and the time when a stimulus was first brought up.

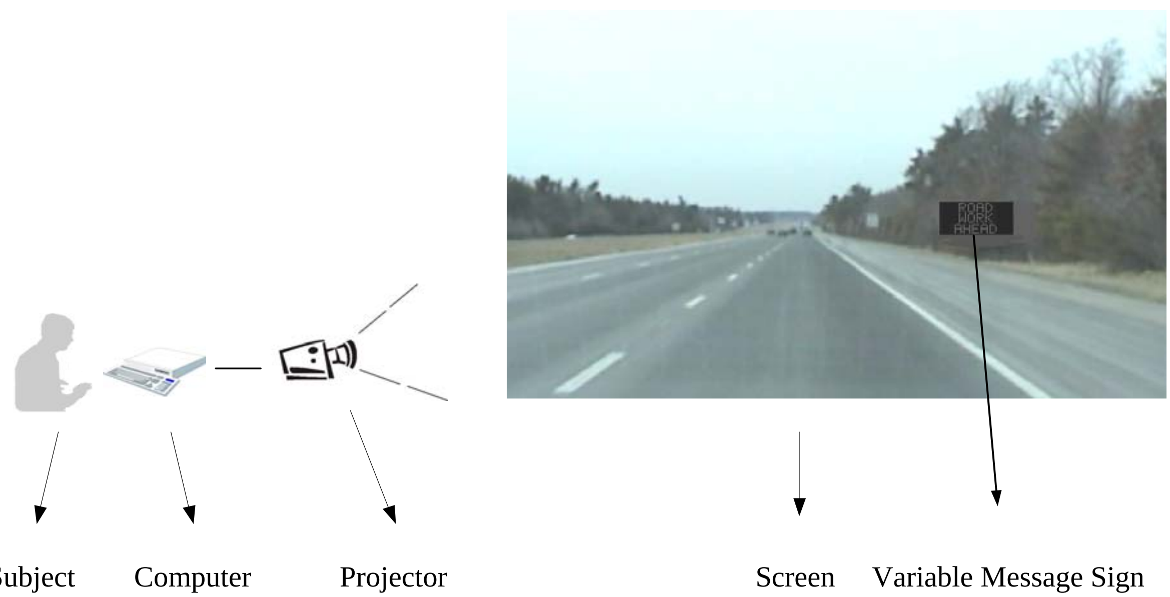


Figure 1 Experiment setup

Eighteen subjects were chosen randomly from the community. Among them, nine were females and nine were males. They were from three age groups, 20~40, 40~60, and above 60. Three females and three males were included in each age group. Each subject was required to have driving experience on interstate highways and with normal or near-normal visions. Before starting the experiment, each subject was screened with a vision test, and was briefed with the purpose and procedures of the experiment. In the beginning, a short warm up session was conducted to familiarize the subjects. The subject needs to enter her/his demographical information through an interactive screen (see Figure 2). If the subject did not make a response properly, e.g., if a response was made too early or too late, an error/warning message would appear on the screen (see Figures 3 and 4) to remind the subject about the mistake she/he just made, and to help prevent it from reoccurring. All of these made sure that every subject understood the experiment and procedures involved before the actual run. With the subject's consent, the actual experiment started.

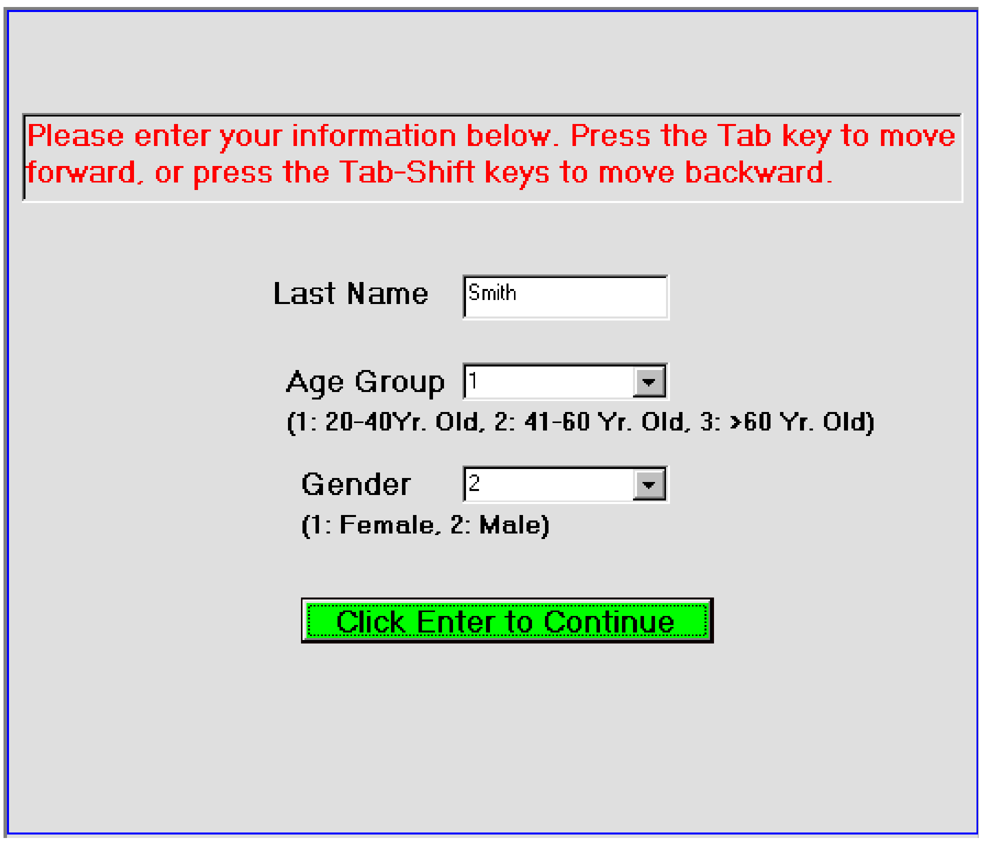


Figure 2. The subject demographical information input window

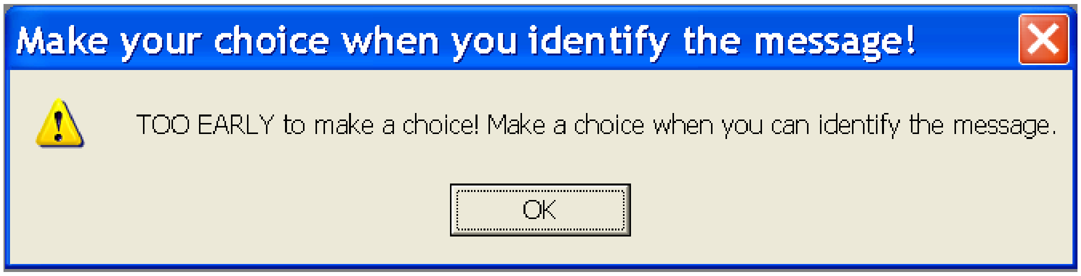


Figure 3. The error message window when a response was made too early

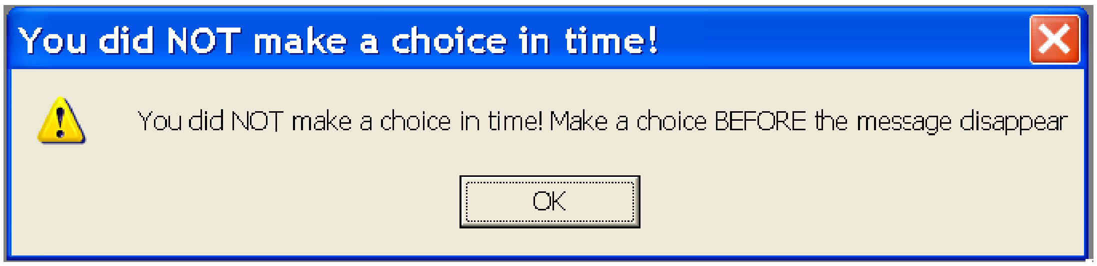


Figure 4. The error message window when a response was made too late

5. PHASE I STUDY

5.1 Experiment Design

In phase I, two groups of factors, including main factors and blocking factors were considered. Controllable factors included font size and font color (foreground color). Blocking factors were drivers' age and gender. A two-factor blocked factorial design was employed to identify the significant effect(s). The corresponding statistical model is shown as follows:

$$T = \mu + C + S + C \times S + A + G + A \times G + \varepsilon \quad (1)$$

where:

T – subjects' response time, second;

μ – overall means, second;

C – font color (main factor);

S – font size (main factor);

A – subjects' age (blocking factor);

G – subjects' gender (blocking factor);

ε – error.

Three different messages, "ACCIDENT AHEAD, USE ALT", "CAUTION, DELAYS AHEAD", and "ROAD WORK AHEAD" were displayed in the type of full matrix (Figure 5). Each VMS stimulus was designed using black as the background color and with a different combination of font color and font size (Table 1). In turn, there were 18 different discretely displayed VMSs in this experiment (3 messages × 3 font colors × 2 font sizes). Since it was possible that a test subject guessed the messages' content by only reading its first word/letter, or by judging its length, 6 additional VMS stimuli with fake messages were added. These messages were composed with minor modifications on the contents of the real messages, i.e., the word(s) or the word order was changed (see Figure 6). Therefore, there were a total of 24 different VMS stimuli in the experiment. These VMS stimuli were merged with a driver's view driving video that was shot on interstate highway 90 westbound in a segment between Worcester and Springfield, MA at about 10 am on a bright sunny day in October 2001. These stimuli were presented to each subject in a random order. Subjects were asked to press key "1" for "ACCIDENT AHEAD, USE ALT", key "2" for "CAUTION, DELAYS AHEAD", key "3" for "ROAD WORK AHEAD", and key "0" for any other messages (fake messages). After a response key was pressed, the current VMS stimulus would disappear and the next stimulus would appear after a random elapse time between 2 and 6 seconds. Each subject went through a total of 72 runs, i.e., 24 different VMS displays with 3 repetitions per VMS display. The response time and accuracy were recorded. In order to collect a subject's impression about the VMS design and display, a short survey was given to each subject after the experiment. The following questions were asked in the survey.

(1) Which is the best font color?

(2) Which is the worst color?

(3) Did you notice the difference in font size in the experiment?

Font Size: 5 × 7 dot

Font Size: 4 × 7 dot

Figure 5. The real messages used in phase I

Table 1. The factors and their levels in phase I

Main Factors	Level
Font color	Green, Red, Yellow
Background color	Black
Font size	5 × 7 dot, 4 × 7 dot
Blocking Factors	Level
Subjects' age	20~40 year, 40~60 year, >60 year
Subjects' gender	Female, Male

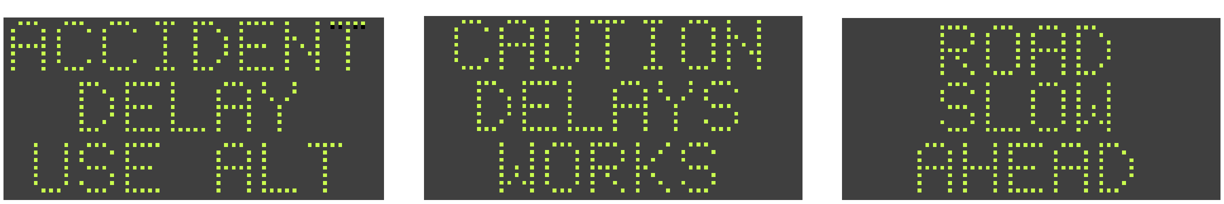


Figure 6. Examples of fake messages used in phase I

5.2 Results and Discussions

Prior to the analysis, the distribution of response times was checked and found to have a mean of 10.645 seconds and a standard deviation of 1.266 seconds. Figure 7 gives a normal plot[3] of all correctly responded data points that were included in the analysis. According to the plot, the data were found slightly deviated from a normal distribution

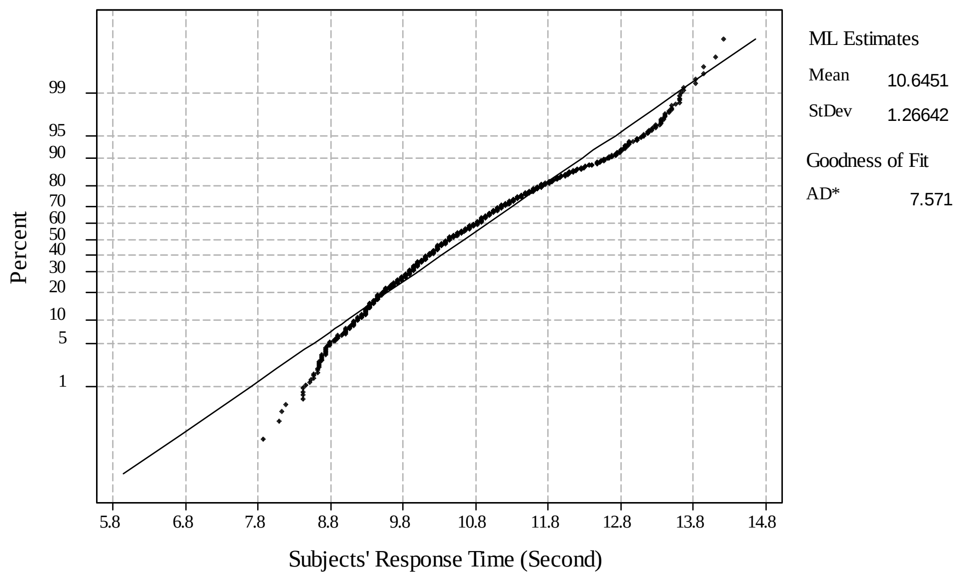


Figure 7. The normal plot of subjects' response time

Using MiniTAB, an analysis of variance (ANOVA) was conducted, and the result is shown in Table 2. From this table, it was found that font color, age, and gender were significant ($p < 0.05$), while other factors and the interactions were not. It shall be noted that a total of 1,296 experiment data were collected but only 833 of them were included in the analysis since the others were discarded due to their incorrect responses. From Figure 8, it can be seen that the

[3] The normal plot is a graphical tool to judge the normality of the distribution of the sample data. The horizontal axis shows the numerical values of the observations, and the vertical axis gives the relative probability frequency of the observed responses. AD stands for Anderson-Darling statistic, which is a criterion for testing the goodness of fit. Usually, the smaller the value, the better the fit.

green font color took the least response time, while the red color took the longest. As to the effects of age on reaction time, the older age group took the least amount of time to respond. This contradicts our common sense. It might be conjectured that older subjects tended to respond faster due to their extensive driving experience. Another possibility could be that older drivers might be more alerted to road traffic and signs than younger drivers. The effect of gender on response time is marginally significant. Female subjects responded a little faster than male subjects.

Table 2. Analysis of variance results

Source	DF	Seq. SS	Adj. SS[4]	Adj. MS	F	P
Font Color	2	70.122	59.018	29.509	23.58	0.000*
Font Size	1	4.938	4.207	4.207	3.36	0.067
Font Color×Font Size	2	0.643	0.919	0.459	0.37	0.693
Age	2	221.201	222.410	111.205	88.88	0.000*
Gender	1	4.247	5.195	5.195	4.15	0.042*
Age×Gender	2	6.315	6.315	3.158	2.52	0.081
Error	822	1028.516	1028.516	1.251		
Total	832	1335.982				

* significance level = 0.05

[4] When the design is not balanced, or when there are missing cells in a full factorial ANOVA design, then there is ambiguity regarding the specific comparisons between the (population, or least-squares) cell means that constitute the main effects and interactions of interest. The adjusted sums of squares (Adj. SS) were required to test "balanced" hypotheses. The Adj. SS for effects can have fewer degrees of freedom than they would have if there were no missing cells, and for some missing cell designs, can even have zero degrees of freedom. The philosophy of Adj. SS is to test as much as possible of the original hypothesis given the observed cells. If the pattern of missing cells is such that no part of the original hypothesis can be tested, so be it.

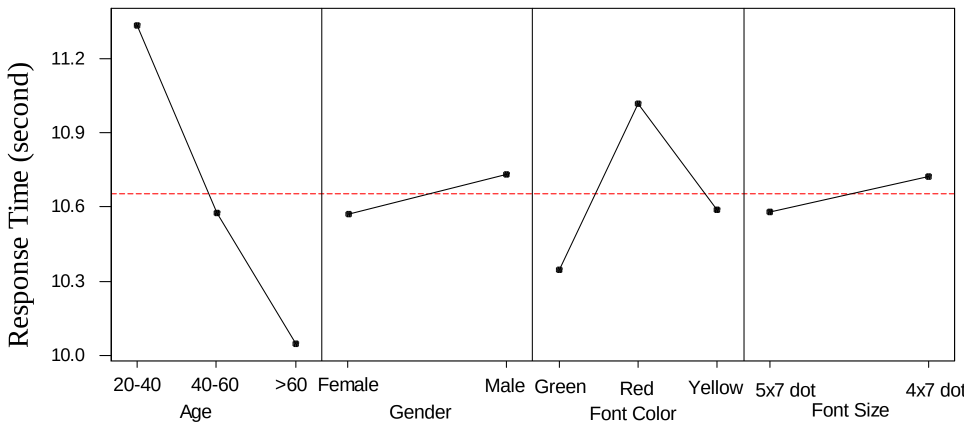


Figure 8. Main effect plots for the factors

To investigate the difference among subjects, additional ANOVAs were conduct among all subjects, and among subjects within the same gender or age groups, or gender and age combinations. The ANOVA summary regarding all subjects is shown in Table 3. It was found that in addition to font color and font size, different subjects gave different response times. Other ANOVA analyses also revealed that there were significant differences between subjects in their individual subgroups in addition to font color and font size (Table 4). Overall, the result of font color's effect on response time is in accordance with the survey result (Table 5). The short survey revealed that 54.8% of the subjects chose green as the best color, followed by yellow (45.2%), no subject chose red. 80.6% of the subjects thought the red color was the worst, followed by yellow (12.9%), and then by green (6.5%). About 29.0% of the subjects noticed the difference in font size in different VMS stimulus, 64.5% did not, and 6.5% were not sure. The accuracy statistics among the three age groups (see Table 6) and the corresponding ANOVA (see Table 7 and Figure 9) shown that older subjects had the lowest accuracy rate while middle-aged subjects had the highest. It also revealed that male subjects had a slightly higher accuracy rate than female subjects, though female subjects proved to respond faster than male subjects.

Table 3. Analysis of variance results considering subject as a blocking factor

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Font Color	2	70.122	72.357	36.178	47.03	0.000*
Font Size	1	4.938	4.306	4.306	5.60	0.018*
Font Color×Font Size	2	0.643	1.134	0.567	0.74	0.479
Subject	17	637.124	637.124	37.478	48.72	0.000*
Error	810	623.155	623.155	0.769		
Total	832	1335.982				

* significance level = 0.05

Table 4. Analysis of variance results by different subject groups

Group	Significance of Factors			
	Main Factors			Blocking Factor
	Main Effect		Interaction	Subject
	Font Color (C)	Font Size (S)	C × S	
Female (9)	*			*
Male (9)	*	*		*
20-40 Yr. (6)	*			*
40-60 Yr. (6)	*	*		*
> 60 Yr. (6)	*			*
Female 20-40 Yr. (3)	*			*
Female 40-60 Yr. (3)	*			*
Female Over 60 Yr. (3)	*			*
Male 20-40 Yr. (3)		*		*
Male 40-60 Yr. (3)	*			*
Male > 60 Yr. (3)	*			*

* significance level = 0.05

Table 5. The result of short follow-up survey

Question	Option	Percentage (%)
Which is the best font color?	Red	0.0
	Green	54.8
	Yellow	45.2
	Not sure	0.0
Which is the worst font color?	Red	80.6
	Green	6.5
	Yellow	12.9
	Not sure	0.0
Did you notice the difference in font size in the experiment?	Yes	29.0
	No	64.5
	Not sure	6.5

Table 6. Accuracy statistics by age and gender

Subjects		Accuracy (%)
20~40 Years Old	Female	83
	Male	89
40~60 Years Old	Female	94
	Male	96
Over 60 Years Old	Female	74
	Male	80

Table 7. ANOVA of accuracy by age and gender

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age	2	0.10979	0.10979	0.05490	5.17	0.024*
Gender	1	0.00836	0.00836	0.00836	0.79	0.392
Age×Gender	2	0.00082	0.00082	0.00082	0.04	0.962
Error	12	0.12750	0.12750	0.01063		
Total	17	0.24648				

* significance level = 0.05

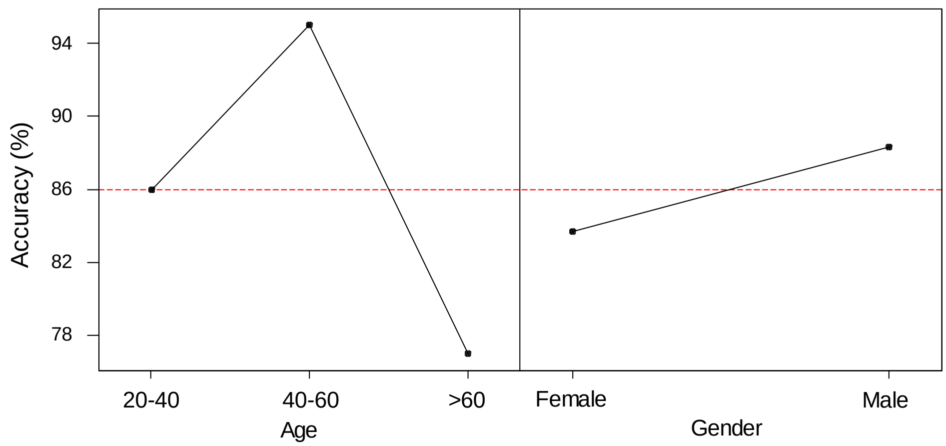


Figure 9. Main effect plot for the factors in accuracy ANOVA

Regression analysis was carried out to investigate if there was a correlation between response time and accuracy. Regression result is shown in Table 8 (a brief description about the regression method employed here is attached in Appendix B). It can be seen that there is no evidence to claim a meaningful correlation between response time and accuracy with an R^2 of 0.016. In other words, accuracy is independent of response time. A breakdown analysis by age and gender groups also yielded no correlation except a marginal correlation found in the 20~40 year age group (Table 9).

Table 8. Overall regression analysis of subjects' average response times on accuracy

Predictor	Coef.	SE Coef.	T	P
Constant	11.486	1.617	7.10	0.000
Accuracy	-0.964	1.869	-0.52	0.613

S = 0.9280 R-Sq = 1.6% R-Sq(adj.) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.2289	0.2289	0.27	0.613
Residual Error	16	13.7775	0.8611		
Total	17	14.0064			

Table 9. Regression analysis of subjects' response time by age and gender

Group	S	R^2	R^2 (Adj.)	Correlation
Female (9)	0.1198	0.6%	0.0%	No
Male (9)	0.1362	5.2%	0.0%	No
20-40 Yr. (6)	0.0687	55.4%	44.2%	Marginal
40-60 Yr. (6)	0.0349	7.0%	0.0%	No
Over 60 Yr. (6)	0.1324	21.3%	1.6%	No

5.3 Conclusion

Based on the above results and discussions, the following conclusions could be reached: Font color, drivers' age, and gender significantly affected response time. Font color "green" was the best among the three colors. Older drivers responded faster than younger drivers but with lower accuracy. Female drivers responded faster than male drivers but also with lower accuracy. No significant correlations were found between subjects' response time and accuracy. The response times obtained between different subjects were significantly different, and this difference was also observed among subjects in the same age group, gender group, and age/gender combination group.

It needs to point out that the above conclusions were based on simulated lab experiments. Validation of these results might require real field studies.

6. PHASE II STUDY

6.1 Experiment Design

Based on the findings from phase I, phase II study focused on the display format of VMSs as viewed by motorists driving under different weather and lane conditions. Similar to phase I, real video scenes were integrated with computer-generated VMSs to create a simulated driving environment. The videos were shot in a segment of I-495 westbound between Rayham and Mansfield, MA at about 10 am in April 2002. VMSs stimuli with different display formats and contents, were introduced in a random but controlled manner. Two groups of factors, main factors and blocking factors, were considered in the experiment (Table 10). Main factors included VMS display format, number of message lines, weather, and driving lane. Blocking factors were drivers' age and gender. Subjects were required to make proper responses to stimulus based on their comprehension. A series of blocked factorial experiments similar to those in Phase I were conducted to fully investigate the effects of these factors and their interactions. Based on the findings from phase I study, each VMS display was designed using black as the background color, green as the font color, and 5 × 7 dot as the font size. To investigate the effects of the main factors, blocking factors and their interactions, a three-factor blocked factorial design was employed. The statistical model in this design is:

$$T = \mu + F + W + L + F\times W + F\times L + W\times L + F\times W\times L + A + G + A\times G + \varepsilon \qquad (2)$$

where:

T – subjects' response time in second;

μ – overall mean;

F – VMS display format;

W – weather;

L – driving lane;

A – subjects' age;

G – subjects' gender;

ε – error.

Table 10. The factors and their levels in the experiment

Main Factors	Level
Font color	Green
Background color	Black
Font size	5 × 7 dot
VMS display	Discrete display, Two-phase sequential display*
Weather	Sunny, Cloudy
Driving lane	Outer, Middle
Blocking Factors	Level
Subjects' age	20-40, 40-60, > 60 year old
Subjects' gender	Female, Male

* message cycle = 3.0 seconds (1.5 seconds/phase, no blank time)

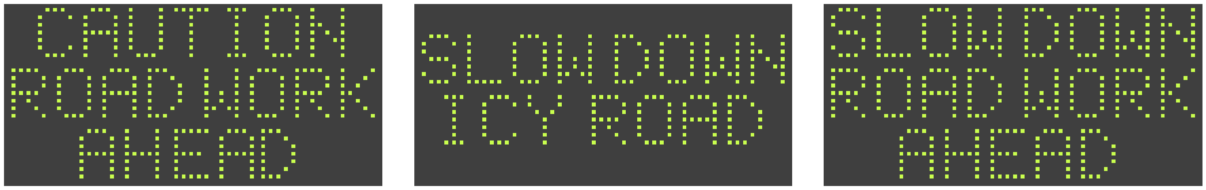
Being a main factor in this experiment, the VMS display format has two settings, a discrete display and a two-phase sequential display. Three one-phase discretely displayed messages and three two-phase sequentially displayed messages were composed in the type of full matrix. Part of the discretely displayed VMS stimuli contained two-line messages and others contained three-line messages (see Figure 10). Among the two-phase sequentially displayed messages, the first frame all contained a single-line message but the second frame contained either a single-line or two-line messages (see Figure 11). The number of message lines was thus considered as another main factor. The weather factor included two settings, sunny and cloudy, and the driving lane factor included driving in the outer lane and in the middle lane. Different settings in the former two factors were generated by the computer. Different conditions in the latter two factors were produced through the driving videos taken with different weather and driving lanes. As a result, there were four weather-driving lane combinations, sunny day and outer lane, sunny day and middle lane, cloudy day and outer lane, and cloudy day and middle lane. Each combination was studied through a session. The experiment setup is the same as that in phase I.

The experiment consisted of four sessions. In each session, six different VMS displays were used (3 messages × 2 display formats). Again, since it was possible that a test subject guessed a message content by only reading its first word/letter in one phase or judging its length, three additional VMS displays with "fake" message contents were added to each session. These messages were composed with minor modifications on the contents from the real messages (see Figure 12). With three repetitions for each VMS display, there were a total of twenty-seven VMS displays for a subject in each session. These VMS displays were presented in a random order. Subjects were required to press key "1" for "CAUTION, ROAD WORK AHEAD", key "2" for "SLOW DOWN, ICY ROAD", key "3" for "SLOW DOWN, ROAD WORK AHEAD", and key "0" for any other messages (fake messages). Each subject went through a total of 108 runs, i.e., 4 sessions with 27 runs per session. The response time and accuracy were recorded (see Table A2 in Appendix A). Similar to phase I, a short survey was given after experiment with the following questions:

(1) Which kind of VMSs display format do you prefer?

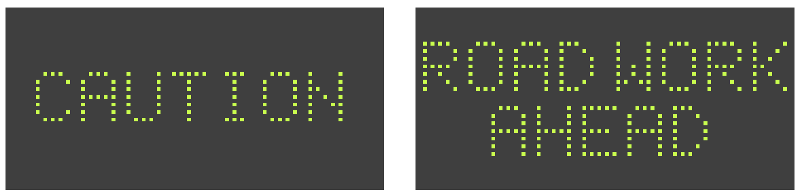
(2) Do you notice the difference in VMSs in different weather?

(3) Which lane do you prefer in order to see the VMSs better?

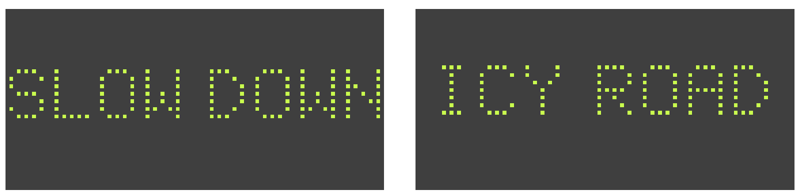


No. 2 No. 4 No. 6

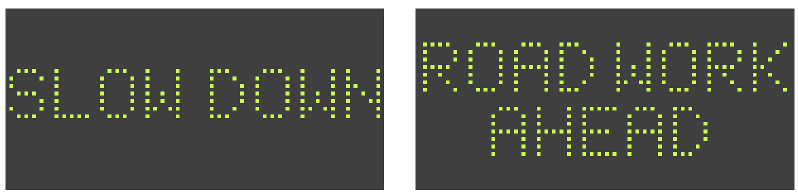
Figure 10. The discrete displayed messages used in the experiment



No. 1



No. 3



No. 5

Figure 11. The two-phase sequential displayed messages used in the experiment

Figure 12. Examples of fake messages used in the experiment

6.2 Results and Discussions

The distribution of these subjects' response times was examined and found to follow a normal distribution with a mean of 6.927 seconds and a standard deviation of 0.856 seconds (Figure 13). It is noted that the average response time in this phase is shorter that that in Phase I, this may be due to the fact that the best font color and font size identified in Phase I were used in Phase II and some learning effects. Also, due to different program modules were used in the two phases, the timer could start differently. However, the relative differences within the same phase remained unchanged. Analysis of variance (ANOVA) was performed, and the result is shown in Table 11. It was found that display format, weather, driving lane, age, and gender were all significant ($p < 0.05$), while the interactions were not. Figure 14 shown that subjects made faster responses to the one-phase discretely displayed VMSs. Faster responses were obtained in a sunny day. It also found that subjects responded faster when driving in the outer lane. The effects of age revealed that younger drivers required the least amount of time to respond, while the middle-aged drivers took the longest. Older drivers responded slightly faster than middle-aged drivers but the difference was insignificant. As to the gender effect, male subjects in general responded faster than female subjects. A breakdown analysis by age and gender was conducted and same results were found with regard to the three main factors (Table 12). The accuracy statistics shown that old group had the lowest accuracy while young group had the highest; and female group had a higher accuracy rate than male group, though male group responded faster than female group (Tables 13, 14 and Figure 15). A correlation analysis did not find enough evidence to support a meaningful correlation between response time and accuracy (Table 15). A breakdown analysis by age and gender groups also yielded no correlation except a marginal correlation found in the 20~40 year age group (Table 16). Generally speaking, the age effects found in this study confirmed the findings from other researches (*1*, *24*, *25*).

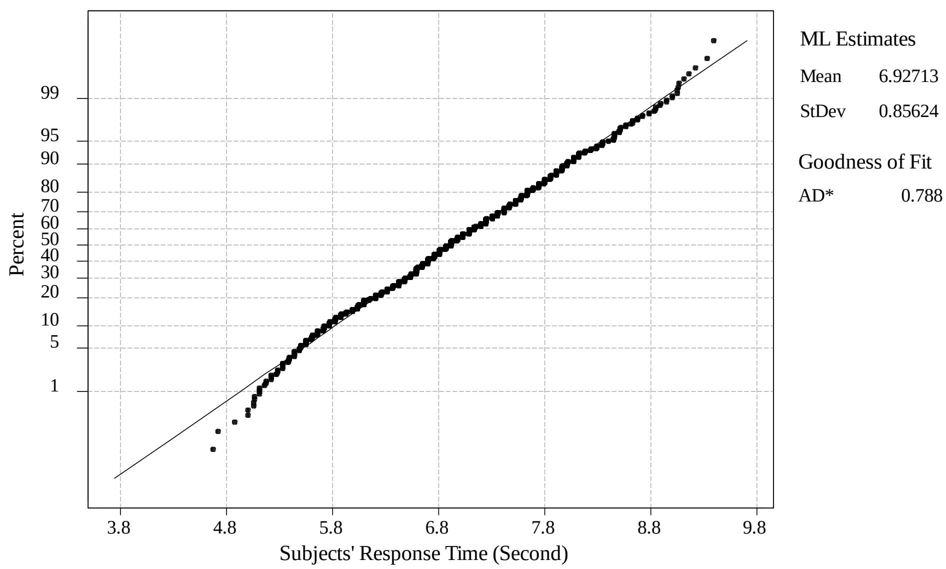


Figure 13. The normal plot of subjects’ response time

Table 11. Analysis of variance results

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Display Format	1	54.449	51.674	51.674	94.01	0.000*
Weather	1	10.827	10.759	10.759	19.57	0.000*
Lane	1	60.704	60.657	60.657	110.35	0.000*
Display Format×Weather	1	0.279	0.270	0.270	0.49	0.483
Display Format×Lane	1	0.029	0.034	0.034	0.06	0.803
Weather×Lane	1	0.162	0.163	0.163	0.30	0.587
Display Format×Weather×Lane	1	0.026	0.026	0.026	0.05	0.828
Age	2	77.372	74.534	37.267	67.80	0.000*
Gender	1	4.229	5.174	5.174	9.41	0.002*
Age×Gender	2	0.712	0.805	0.403	0.73	0.481
Error	1086	596.929	596.929	0.550		
Total	1098	805.719				

* significance level = 0.05

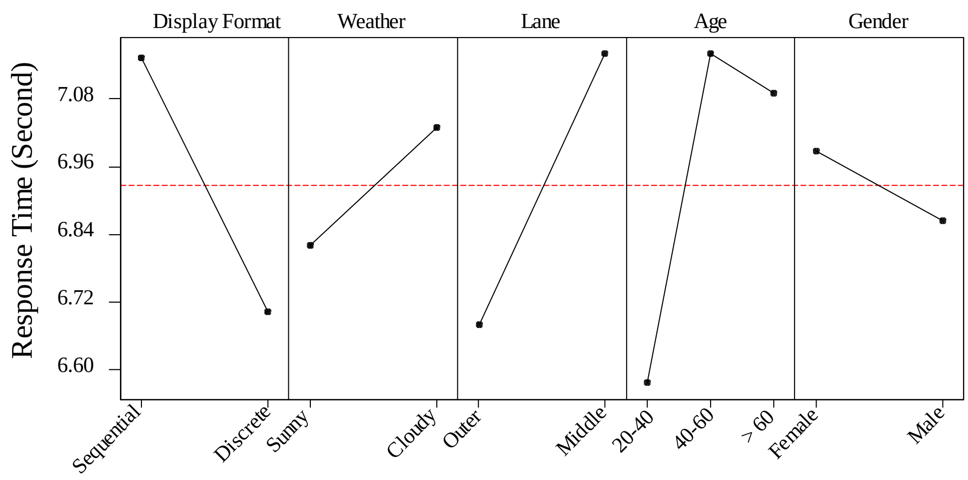


Figure 14. Main effect plots of significant factors

Table 12. Analysis of variance results by age and gender groups

Group	Significance of Factors							
	Main Factors							Blocking Factor
	Main Effect			Interaction				Subject
	Display Format (F)	Weather (W)	Lane (L)	F×W	F×L	W×L	F×W×L	
Female (9)	*	*	*					*
Male (9)	*	*	*					*
20-40 Yr. (6)	*	*	*					*
40-60 Yr. (6)	*	*	*					*
> 60 Yr. (6)	*	*	*					*
Female 20-40 Yr. (3)	*	*	*					*
Female 40-60 Yr. (3)	*	*	*					*
Female Over 60 Yr. (3)	*		*			*		*
Male 20-40 Yr. (3)	*		*			*		*
Male 40-60 Yr. (3)	*		*					*
Male > 60 Yr. (3)	*		*					*

* significance level = 0.05

Table 13. Accuracy by age and gender

Subjects		Accuracy (%)
20-40 Years Old	Female	94
	Male	91
40-60 Years Old	Female	87
	Male	85
Over 60 Years Old	Female	77
	Male	74

Table 14. ANOVA of accuracy by age and gender

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Age	2	0.087048	0.087048	0.043524	7.01	0.010*
Gender	1	0.002427	0.002427	0.002427	0.39	0.544
Age×Gender	2	0.000061	0.000061	0.000030	0.00	0.995
Error	12	0.074513	0.074513	0.006209		
Total	17	0.164049				

* significance level = 0.05

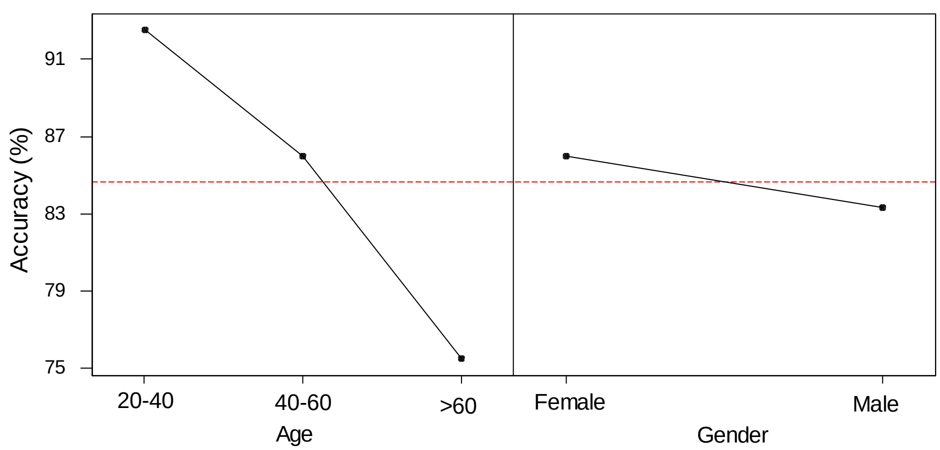


Figure 15. Main effect plot for the factors in accuracy ANOVA

Table 15. Overall regression analysis of subjects' average response times on accuracy

Predictor	Coef	SE Coef	T	P
Constant	1.2313	0.3690	3.34	0.004
Average	-0.05563	0.05306	-1.05	0.310

S = 0.09921 R-Sq = 6.4% R-Sq(adj) = 0.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.010819	0.010819	1.10	0.310
Residual Error	16	0.157490	0.009843		
Total	17	0.168310			

Table 16. Regression analysis of subjects' response time by age and gender

Group	S	R^2	R^2 (Adj.)	Correlation
Female (9)	0.525757	8.1%	0.0%	No
Male (9)	0.417467	7.5%	0.0%	No
20-40 Yr. (6)	0.216971	68.9%	61.2%	Marginal
40-60 Yr. (6)	0.574690	0.7%	0.0%	No
Over 60 Yr. (6)	0.263839	0.1%	0.0%	No

To analyze the effect of number of message lines, two additional ANOVAs were conducted on the discretely displayed VMSs and the sequentially displayed VMSs respectively (Tables 17 and 18). In each group, number of message lines (NML), weather, driving lane, age, and gender all exhibited significant effects (Figures 16 and 17). For the discretely displayed VMSs, messages number 2 and 6 were three-line VMSs, while number 4 is a two-line VMS. Figure 16 shown that the two-line message sign (number 4) took less response time than the three-line message signs (numbers 2 and 6). As to the sequentially displayed VMSs, messages number 1 and 5 were two-phase VMSs in which the first phase had a single-line message and the second phase had a two-line message while message number 3 was a two-phase VMS in which both phases used single-line message. From Figure 17, it found that message 3 took less response time than messages 1 and 5. In other words, a single line message demands less processing and results in less response time than a multiple line message. These are in consistence with the findings of Miller et al (*21*) and confirm that the simple and concise message components result in quick responses.

Table 17. Analysis of variance results the discretely displayed VMSs

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Lane	1	31.8129	32.0700	32.0700	62.69	0.000*
Weather	1	3.8130	5.0320	5.0320	9.84	0.002*
NML	2	49.1492	49.1958	24.5979	48.09	0.000*
Lane×Weather	1	0.0109	0.0088	0.0088	0.02	0.896
Lane×NML	2	0.0411	0.0374	0.0187	0.04	0.964
Weather×NML	2	0.3744	0.3646	0.1823	0.36	0.700
Lane×Weather×NML	2	1.0899	1.0899	0.5450	1.07	0.345
Age	2	42.2553	45.5295	22.7648	44.50	0.000*
Gender	1	1.6604	2.0206	2.0206	3.95	0.047*
Age×Gender	2	0.2054	0.2362	0.1181	0.23	0.794
Error	536	274.1815	274.1815	0.5115		
Total	552	404.5940				

* significance level = 0.05

Table 18. Analysis of variance results the sequentially displayed VMSs

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Lane	1	29.0221	25.1064	25.1064	60.43	0.000*
Weather	1	7.2560	6.8915	6.8915	16.59	0.000*
NML	2	47.4337	46.9504	23.4752	56.50	0.000*
Lane×Weather	1	0.1792	0.2076	0.2076	0.50	0.480
Lane×NML	2	1.8727	1.8577	0.9289	2.24	0.108
Weather×NML	2	1.4728	1.4466	0.7233	1.74	0.176
Lane×Weather×NML	2	0.0862	0.0862	0.0431	0.10	0.901
Age	2	34.4245	35.9543	17.9772	43.27	0.000*
Gender	1	2.6209	3.3126	3.3126	7.97	0.005*
Age×Gender	2	1.4673	1.5186	0.7593	1.83	0.162
Error	529	219.7834	219.7834	0.4155		
Total	545	345.6190				

* significance level = 0.05

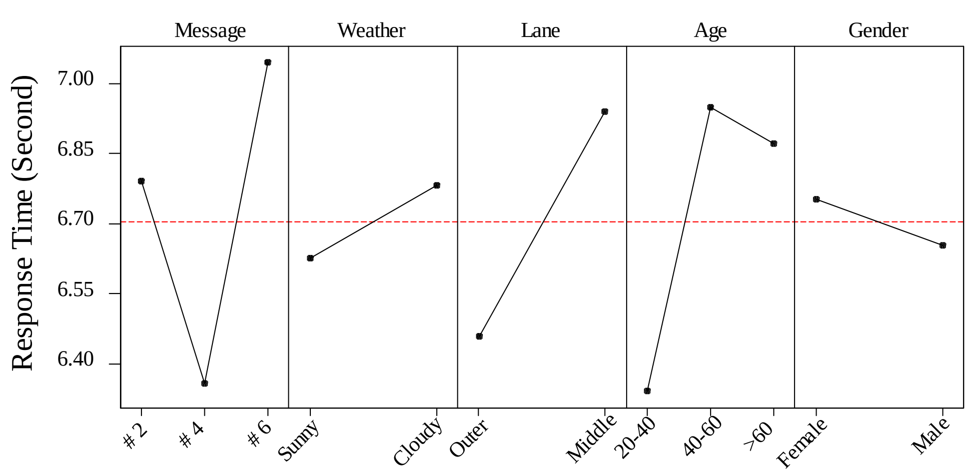


Figure 16. Main effect plots of significant factors for the discretely displayed VMSs

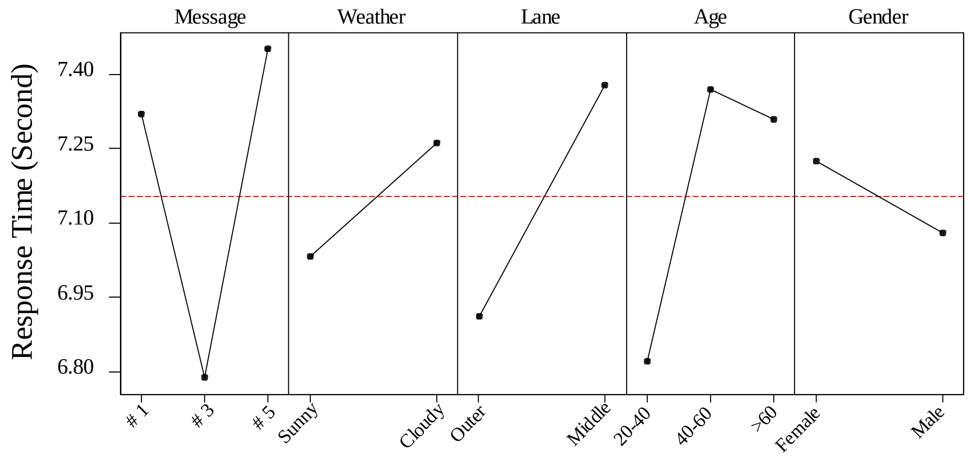


Figure 17. Main effect plots of significant factors for the sequentially displayed VMSs

To investigate the difference among subjects, eighteen more ANOVAs were conducted on each subject (Table 19). These analyses revealed that there were some differences among different subjects. In 66.7% of the subjects, display format was significant; in 27.8%, weather was significant; and in 77.8%, driving lane was significant. Overall, the statistical result was consistent with the survey result. The survey revealed that (Table 20), 71.4% of the subjects preferred the discrete display format because most of them thought "…VMSs displayed in sequence are difficult to identify…", or "...sometimes first saw the second part of a sequential message…had to wait until seeing the whole cycle..", while 14.3% preferred the sequential display format because of "…short and concise information in each phase for easy scanning…" or "…sometimes the discrete displayed information are overwhelming…", 11.4% thought there was no difference between these two, and only 2.9% were not sure. As to the weather influence, 28.6% of the subjects noticed the differences in VMSs under different weather, 51.4% did not because they "concentrated on identifying the message contents and making responses…", and

20% were not sure. For the lane preference, 54.3% chose the outer lane, 11.4% chose the middle lane, and 34.3% felt no difference at all.

Table 19. Individual analysis of variance results by subject

Subject	Age	Gender	Significance of Main Factors						
			Main Effect			Interaction			
			Format (F)	Weather (W)	Lane (L)	F×W	F×L	W×L	F×W×L
1	20-40	Female	*		*			*	
2	20-40	Female	*	*	*				
3	20-40	Female	*	*	*				
4	20-40	Male			*			*	
5	20-40	Male	*		*				
6	20-40	Male	*		*		*	*	
7	40-60	Female	*	*	*				
8	40-60	Female	*						
9	40-60	Female	*	*					
10	40-60	Male			*				
11	40-60	Male	*	*					
12	40-60	Male	*		*				
13	> 60	Female			*				
14	> 60	Female	*		*				
15	> 60	Female	*		*			*	
16	> 60	Male			*				
17	> 60	Male	*						
18	> 60	Male			*				

* significant in the level = 0.05

Table 20. Short follow-up survey results

Question	Option	Percentage (%)
Which kind of VMSs display format do you prefer?	Discrete display	71.4
	Sequential display	14.3
	Both/No difference	11.4
	Not sure	2.9
Do you notice the difference in VMSs in different weather?	Yes	28.6
	No	51.4
	Not sure	20.0
Which lane do you prefer in order to see the VMSs better?	Outer lane	54.3
	Middle lane	11.4
	No difference	34.3
	Not sure	0.0

6.3 Conclusions

Based on the above results and discussions, the following conclusions might be reached:

All experiment factors considered in the experiment significantly affect response time, they are: VMSs display format, number of message lines, weather, and driving lane. No interaction among them was found significant. Regarding the display format, a discretely displayed message took less response time than a sequentially displayed message. As to the number of message lines, less is better, a single-line message is better than a multiple-line message. When driving in the outer lane, motorists can better view and respond to a VMS than driving in the middle lane. Driving at highway speed, a VMS can be better seen in a sunny day than a cloudy day. Also, it found that older drivers exhibit slower response and less accuracy than younger drivers; female drivers exhibit slower response but higher accuracy than male drivers.

Again, the above conclusions were based on simulated lab experiments. Validation of these results might require real field studies.

7. RECOMMENDATIONS

Based on the result and analysis of this study, the following recommendations might be stated:

Both 4×7 dot and 5×7 dot could be used as the font size in VMS messages, though font size 5×7 dot is a little better than 4×7 dot; of the three font colors (green, red, and yellow), green might be the best font color, followed by yellow and red; a discretely displayed message took less response time than a sequentially displayed message; simple and concise VMS message design is recommended.

It should be noted that the above recommendation was reached through a simulated highway-driving experiment, therefore some measures were adapted to lab settings. Under real dynamic driving on highways, it should consider the reading time (the time that actually takes a driver to read a VMS) rather than the response time (the elapsed time between the VMS's first exposure and the response) in the analysis. The conspicuity of VMS, that is, the ability of a VMS to stand out clearly in "visual noises/distractions" (the complex backgrounds that adversely affect sign detection and legibility, such as overpasses, buildings, trees, advertising signs, vehicles, etc.) was not considered here. The test subjects although situated in a "virtual" driving environment, were not responsible of actual driving, and thus could focus entirely on the VMS display. With the advance of technology, it is hoped that more realistic study in this area under the premise of economic and reliable means will be conducted in the future.

8. REFERENCES

1. Dudek, C. L. *Guidelines on the Use of Changeable Message Signs.* Report FHWA-TS-90-043. FHWA, July 1991.
2. Benson, B. G. Motorist Attitudes about Content of Variable Message Signs. *Transportation Research Record 1550*, 1996, pp. 48-57.
3. MV2. Evaluation de La Politique D'affichage Sur Les Panneaux a Message Variable. Report to Direction Regionale de L'Equipment. 1997.
4. Rämä, P. Effects of Weather-controlled Variable Speed Limits and Warning Signs on Driver Behavior. *Transportation Research Record 1689*, 1999, pp. 53-59.
5. Yim, Y., and J. L. Ygnance. Link Flow Evaluation Using Loop Detector Data: Traveler Response to Variable Message Signs. *Transportation Research Record 1550*, 1996, pp. 58-64.
6. Wardman, M. et al. Driver Response to Variable Message Signs: A Stated Preference Investigation. *Transportation Research Part C*, Vol. 5, Issue 6, 1997, pp. 389-405.
7. Chaterjee, K. et al. Driver Response to Variable Message Sign Information in London. *Transportation Research Part C*, Vol. 10, Issue 2, 2002, pp. 149-169.
8. Emmerink, R. H. M. et al. Variable Message Signs and Radio Traffic Information: An Integrated Empirical Analysis of Drivers' Route Choice Behavior. *Transportation Research Part A*, Vol. 30, Issue 2, 1996, pp. 135-153.
9. Rämä, P., and R. Kulmala. Effects of Variable Message Signs for Slippery Road Conditions on Driving Speed and Headways. *Transportation Research Part F*, Vol. 3, Issue 2, 2000, pp. 85-94.

10. Luoma, J. et al. Effects of Variable Message Signs for Slippery Road Conditions on Reported Driver Behavior. *Transportation Research Part F*, Vol. 3, Issue 2, 2000, pp. 75-84.

11. Comte, S. L., and A. H. Jamson. Traditional and Innovative Speed-reducing Measures for Curves: An Investigation of Driver Behavior Using a Driving Simulator. *Safety Science*, Vol. 36, Issue 3, 2000, pp. 137-150.

12. Garber, N. J., and S. Srinivasan. *Effectiveness of Changeable Message Signs in Controlling Vehicles in Work Zones*. Report VTRC 98-R10. Virginia Transportation Research Council, Charlottesville, Virginia, December 1998.

13. *Variable Message Signs*. http://www.path.berkeley.edu/~leap/travelerinfo/Driver_Info/message.html. Accessed Oct. 6, 2001.

14. Kraan, M. et al. Evaluating Network Wide Effects of VMSs in the Netherlands. *The 78th Annual Meeting of the TRB*, Washington D.C., January 1999.

15. Metaxatos, P., and S. Soot. Evaluation of Driver'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.

16. Dudek, C. L. Changeable Message Sign Messages for Work Zones. *Transportation Research Record 1692*, 1999, pp. 1-8.

17. Durkop, B. R., and C. L. Dudek. Texas Driver Understanding of Abbreviations for Changeable Message Signs. *The 80th Annual Meeting of the TRB*, Washington D.C., January 2001.

18. Hustad, M. W., and C. L. Dudek. Driver Understanding of Abbreviations for Changeable Message Signs in New Jersey. *The 78th Annual Meeting of the TRB*, Washington D.C., January 1999.

19. National Cooperative Highway Research Program Synthesizes Research on Changeable Message Signs. http://www.usroads.com/journals/p/rej/9709/re970903.htm. Accessed June 6, 2001.

20. Garvey, P. M., and D. J. Mace. *Changeable Message Sign Visibility*. Report FHWA-RD-94-077. FHWA, April 1996.

21. 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.

22. Miller et al. Effective Use of Variable Message Signs: Lessons Learned Through Development of Users' Manuals. *The 74th Annual Meeting of the TRB*, Washington D.C., January 1995.

23. Dudek, C. L., and G. L. Ullman. Flashing Messages, Flashing Lines, and Alternating One Line on Changeable Message Signs. *The 81st Annual Meeting of the TRB*, Washington D.C., January 2002.

24. Anttila, V. et al. Visual Demand of Bilingual Message Signs Displaying Alternating Text Messages. *Transportation Research Part F*, Vol. 3, Issue 2, 2000, pp. 65-74.

25. Dewar et al. Age Differences in the Comprehension of Traffic Sign Symbols. *The 73rd Annual Meeting of the TRB*, Washington D.C., January 1994.

APPENDIX A: EXPERIMENT DATA

Table A1. Phase I experiment data example

(Please see the whole phase I data in file named “VMS1Data.xls” in CD)

Subject	Age	Gender	VMS	Response Time (second)	Accuracy
St. Vincent	1	1	11	10.10938	0
Shalhoub	1	2	18	9.95312	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 “Accuracy”, 0 stands for the incorrect response to a real VMS message, 1 stands for the correct response to a real VMS message, 2 stands for the correct response to a fake VMS message, which means subjects successfully identified the message was a fake one, and 3 stands for the incorrect response to a fake VMS message, which means that subjects could not identify it was a fake message.

Table A2. Phase II experiment data example

(Please see the whole phase II data in file named “VMS2Data.xls” in CD)

Subject	Age	Gender	VMS Format	Lane	Weather	Response Time (second)	VMS	Accuracy
Thompson	1	1	1	1	2	7.361328	1	1
Field	1	2	2	1	1	7.199219	4	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 “VMS Format”, 1 stands for sequential display, and 2 stands for discrete display.

4. In column “Lane”, 1 stands for outer lane, and 2 stands for middle lane.

5. In column “Weather”, 1 stands for sunny days, and 2 stands for cloudy days.

6. In column “Accuracy”, 0 stands for the incorrect response to a real VMS message, 1 stands for the correct response to a real VMS message, 2 stands for the correct response to a fake VMS message, which means subjects successfully identified the message was a fake one, and 3 stands for the incorrect response to a fake VMS message, which means that subjects could not identify it was a fake message.

APPENDIX B: REGRESSION ANALYSIS

The Regression Equation

A linear regression equation is simply a mathematical equation for a line. It is the equation that describes the regression line. In algebra, we represent the equation for a line with something like this:

$$y = a + bx$$

a is the intercept, or the point at which the line travels through the y-axis (sometimes called the y-intercept), and b is the slope of the line. One can think of the y-intercept as the value of y when x is equal to 0. With a grid, we could find the slope of the line by counting how many points we have to go up to meet the line again after we have gone over one point to the right (remember "rise over run"). So the slope is a ratio of the increase in y with every point increase in x. With regression analysis, we need to find out what the equation of the line is for the best fitting line. What is the slope and intercept for the regression line? If the slope is zero, there is no relationship between x and y. If the slope is larger than 0 (or smaller, if the relationship is negative), there is a relationship.

To figure out the equation for the regression line, we first want figure out the slope, b. Here is the formula for that:

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

Notice that on the top of the formula, we compute the deviations of the x's from the mean of x and the deviation of the y's from the mean and multiple them. We do not square them. This top part of the equation can be called the <u>covariance</u> of x and y. The slope then represents the amount that x and y co-vary together relative to the overall variation of x. And sometimes the equation for b is written as:

$$b = \frac{Cov(xy)}{Var(x)}$$

the intercept, a, can then be obtained using b:

$$a = \bar{y} - b\bar{x}$$

where $\bar{x}$ and $\bar{y}$ are the means of x and y respectively. In regression analysis, we are attempting to predict y based on x scores, so we represent the regression equation with a $\hat{y}$ symbol to indicate a predicted score:

$$\hat{y} = a + bx$$

R – Square and Adjusted R - Square

R-square, often called the coefficient of determination, is defined as the ratio of the sum of squares explained by a regression model and the "total" sum of squares around the mean

$$R^2 = 1 - SSE / SST$$

in the usual ANOVA notation. Most people refer to it as the proportion of variation explained by the model, but sometimes it is called the proportion of variance explained. This is misleading because SST is not the variance of Y. In sample terminology, variances are "mean squares." Thus the estimated variance of Y is MST = SST/(n-1) and the estimated residual or error variance is MSE = SSE/(n-p-1) where p is the number of predictors in the regression equation. We "average" by dividing by degrees of freedom rather than by n in order to make the sample mean squares unbiased estimates of the population variances.

The adjusted R square is the R-square weighted by the number of independent variables and observations. Adjusted R-square = R-square - [(k - 1)/(t-k)] * (1-R-square), where k is the number of independent variables and t is the total number of observations.

The R square can never decrease and must either increase or at least stay the same if we add more variables. However, we do not want to add too many variables that do not seem to contribute much to the model, and the adjusted R square tries to take this into account.

The adjusted R square has the advantage over the normal R square that it will not always increase when we add variables, but only increases if variables add significantly to the model. However, it no longer has exactly the same interpretation as the R-square in the sense of share of "variance explained". It may thus be better to use an F test for the significance of the model as a measure of fit.