Using NASA-Task Load Index to Assess Drivers' Workload on Freeway Guide Sign Structures

Fengxiang QIAO[1], Rong ZHANG[2] and Lei YU[3]

[1] Department of Transportation studies, Texas Southern University, 3100 Cleburne Avenue, Houston, TX 77004; PH (713) 313-1915; FAX (713) 313-1856; email: qiao_fg@tsu.edu

[2] Department of Transportation studies, Texas Southern University, 3100 Cleburne Avenue, Houston, TX 77004; PH (713) 313-7532; FAX (713) 313-1856; email: samedolt@gmail.com

[3] Department of Transportation studies, Texas Southern University, 3100 Cleburne Avenue, Houston, TX 77004; PH (713) 313-7282; FAX (713) 313-1856; email: yu_lx@tsu.edu

ABSTRACT

Current research and practices on guide signs studies mainly focus on how to set up a particular guide sign. Less attention is on the tolerable or optimal number of advance guide signs on a sign structure. This paper will assess drivers' workload on various freeway guide sign structures and sign information units, and propose the desirable number of guide sign panels and symbol signs on one freeway guide sign structure. A posterior questionnaire survey was conducted after drivers finished the driving simulator test. Drivers' perception and cognition on different sign structures are carefully analyzed using the NASA-Task Load Index. With the statistical analyses of questionnaires and the recorded data from driving simulator tests, the optimal advance placements of guide signs considering drivers' workload are proposed. Recommendations are provided on how to better deliver necessary information to motorists under various geometric and traffic conditions.

INTRODUCTION

Driving on highways is part of daily lives in modern metropolitan areas. Efficient driving while keeping safety becomes very important, which attracts more and more attentions in the area of transportation studies. It is highly possible that improper driving behavior may cause unexpected incidents so that other vehicles in vicinity might be delayed and stuck in traffic, causing non-recurring congestions and extra exhaust pollutions affecting the air quality. Improper driving behaviors include the situation when drivers are distracted by the overloaded information being

acquired during driving with too much to handle at a time.

The term "workload" is normally used to measure the amount of work a driver should conduct en route. There are three major factors that will affect the drivers' workload. (a) Driver state affecting factors such as fatigue, monotony, sedative drugs, alcohol, and so on. One of the most critical factors known to impact driver performance is fatigue, which may be implicated in around 15% of all fatal large-truck-related crashes (Pratt, 2003). (b) Environmental factors, which may include road environment demands, traffic demands, vehicle ergonomics, automation, and feedback. Research has demonstrated that monotony of road environment can contribute to fatigue and have adverse effects on driving performance (Thiffault and Bergeron, 2003). (c) Driver trait factors such as experience, age, strategy, and so on. A research find information technologies that provide distractions to the driving task may be dangerous for field dependent drivers with low working memories (Lottridge and Chignell, 2007). Freeway guide signs are one of the environmental factors that affect drivers' workloads (Qiao, et, al. 2008). The questions are: (a) How much do freeway guide signs affect drivers' workloads? and (b) Is there any significant difference of drivers' workload on different freeway guide sign structures? This paper is to identify the framework and appropriate number of freeway guide signs on any single guide sign structure. The key hypothesis is "The placement of guide signs affects drivers' workloads and behaviors."

STATE OF THE ART

Traffic guide signs will help drivers to their destinations effectively and safely. Pline (1992) believes that traffic signs are more effective when they comply with the following requirements: (1) fulfill a need command attention, (2) convey a clear and simple message, (3) command respect of the road users, and (4) give adequate time for proper response. The properly placed guide signs may also keep drivers from fatigue and enhance the safety. The study by Thiffault and Bergeron (2003) revealed that monotony of road environment can contribute to fatigue and have adverse effects on drivers' performance. Other research also suggests that reduced workload may not always be beneficial to performance (Funke et. al, 2007).

On the other hand, although human errors are blamed for nearly 90% of traffic crashes (Clark and Funkhouser, 2007), some researchers such as Clark and Funkhouser (2007) believe that traffic incidents may not always be the fault of drivers. A research conducted by Cai *et al*. (2007) found that compared with properly placed signs, improperly placed signs are less effective and cause more stress to the drivers because defective signs may require more visual resources and drivers need to make a huge effort to decipher them. The effectiveness of defective

signs is compromised because drivers may miss them or have insufficient time to make a correct response. Lottridge and Chignell (2007) thought information technologies that provide distractions to the driving task may be dangerous for field drivers with low working memories. Another research indicated that drivers overloaded with information from in-vehicle systems significantly increased the chance of vehicle crashes (Wu, 2007). It is obvious that swarming signs with overloaded information will increase drivers' workload and may cause inappropriate reactions or even result in safety problems. Therefore, it is very important to assess drivers' workload on different sign structures which momentarily demand drivers' attention at the same time.

Assumed as a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance, workload is defined as human-centered, rather than task-centered (Hancock, et. al, 1988). It is experienced as a natural consequence of many daily activities. However, a formal requirement to quantify such an experience using experimentally-imposed rating scales is not a natural or commonplace activity and may result in qualitatively different responses. For this reason, Turksen and Moray (Turkson Et. al, 1985) suggested that the less precise "linguistic" approach provided by fuzzy logic might be appropriate for workload measurement because people naturally described their experiences with verbal terms and modifiers (e.g., "high", "easy", or "moderate") rather than with numerical values. Wood et al. (Wood et. al, 2004) discussed the relationship between different topic types and drivers' workload during driving performance. They believed that using driver's speech is a possible new tool in the arsenal of workload measures especially while driving. Reimer et. al. (2006) concluded that older drivers appeared to have more difficulty acclimating to the simulated driving environment by collecting of heart rate, a traditional physiological measure, which provided additional insight into the differential workload impact on younger and older drivers. Recently, The Peripheral Detection Task (PDT) has been explored as a promising method to measure drivers' workload (Jahn et. al, 2005), and it has been suggested as a standard method to assess safety-relevant workload from the use of in-vehicle information systems (IVIS) while driving.

Among abound workload assessment techniques, subjective ratings such as NASA-TLX are the most commonly used method. Hart (2007) defined NASA-TLX as a multi-dimensional scale designed to obtain workload estimates from one or more operators while they are performing a task or immediately afterwards. Though it was initially designed for use in aviation (Hart, 2007), NASA-TLX was also widely used in transportation studies. Since both laboratory and real-world tasks have the same basic human activities in common, laboratory research results and well-designed measurement tools can be applied a across domains (Hat, 2007).

Comparing other workload assessment methods (Hancock et. al, 1988), subjective ratings may come closest to the essence of mental workload and provide the most generally valid and sensitive indicator. That's why NASA-TLX is selected to evaluate drivers' workload in this study.

REVIEW ON STATE OF THE PRACTICE

Based on the review of relevant existing national and state manuals and research papers, it is found that current research and practice mainly focus on how to set up a particular guide sign (MTSM; TxDOT, 2006; MTH; MnDOT, 2000; FHWA, 2009; MTHBC, 2000; MoDOT, 2006) with less attention on the tolerable or optimal number of freeway guide signs under free flow traffic. The current status of freeway guide sign placement is summarized in the following parts. (1) The US national MUTCD does provide guidelines for guide sign placement regulating the sign's size, color, position, and so on (MTSH; FHWA, 2009). (2) Most states in U.S. follow national MUTCD and American Association of State Highway and Transportation Officials (AASHO) (TxDOT, 2006) Guidelines when placing guide signs (MTH; MnDOT, 2000; FHWA, 2009; MTHBC, 2000; MoDOT, 2006). (3) There are very limited provisions on how to determine the suitable number of guide signs on freeway (TxDOT, 2006; FHWA, 2009). (4) MUTCD stipulates "Advance Guide signs should be placed at 1 km or 0.5 miles and at 2 km or 1 mile in advance of the exit with a third Advance Guide sign placed at 4 km (2 mi) in advance of the exit if spacing permits." which has very limited information about the permitted number of freeway guide sign for each single destination. (5) MUTCD limits the number of different guide signs for different interchanges, such as "At minor interchanges, only one Advance Guide sign should be used." and "No more than one Supplemental Guide sign should be used on each interchange approach." (MUTCD, 2E). (6) Very limited research papers have been found to discuss the impacts of drivers' response and cognition on advance guide sign placement.

The research objectives are: (1) to assess drivers' workload on different guide sign structures, (2) to find a way of appropriately providing guidance to drivers when the total number of signs is limited, and (3) to provide recommendations on identifying the suitable number of freeway guide signs on each guide sign structure.
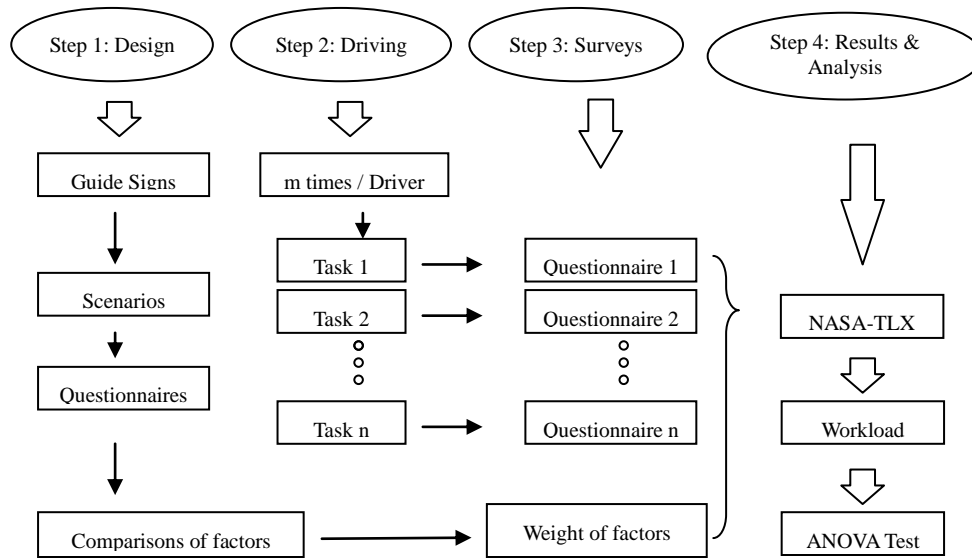
FRAMEWORK OF WORKLOAD TESTING

The four-step framework of the workload testing is illustrated in Figure 1.

Step 1: Design

This includes the design of (1) freeway guide signs, freeway guide sign structures; (2) scenarios design in HyperDrive; and (3) posterior questionnaires.

MUTCD distinguishes the "Guide Sign Information Units" from every single word, number or symbol. Words which present a complete, independent meaning are considered as one guide sign information unit, such as "Polk Blvd", "1 Mile", and "Exit only". Meanwhile, single word, number or symbol such as "EXIT", "44" or the black arrow which are giving specific meaning are also counted as guide sign information units, shown as in Figure 2. SignCAD, a powerful sign design software complying sign design guidelines of the nation and all states, was used to prepare all freeway guide signs in this study.

Figure 1. Framework of workload testing
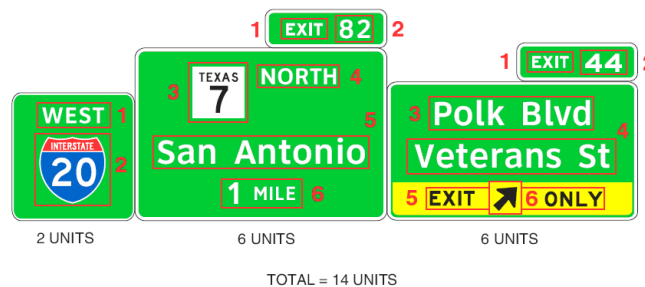
Figure 2. Example of guide sign information units (source: 1)

MUTCD limits the maximum amount of information on any freeway guide sign structure, and there should be no more than 20 information units in total for each freeway guide sign structure. Table 1 lists the desirable and maximum amount of information per freeway guide sign structure according to different number of sign panels. Scenarios were designed using the software HpyerDriver, which permits

customized driving and environmental models in a drag and dropping assembly model. HyperDrive allows the definition of traffic conditions, environmental states, and scripted events during the simulated drive. Freeway guide signs from SignCAD were placed on different guide sign structures associated with various scenarios.

Table 1. Maximum Amount of Information per Freeway Guide Sign Structure

Numbers of Sign Panels	Condition	Maximum Unit of information per structure
2	Desirable	12
	Maximum	16
3	Desirable	16
	Maximum	18
4	Desirable	18
	Maximum	20
5	Desirable	Undesirable Design
	Maximum	20

The posterior questionnaires were designed to evaluate the drivers' workload. The previously mentioned NASA-TLX, in the form of weights, can be applied to ratings for specific factors. Two set of questionnaires were designed based on the requirement of NASA-TLX. The first one was to collect self-rating scores for the six factors which were indicated as NASA-TLX. This set of questionnaires is almost the same for all subjects except for the first question which involved different guide sign structures. The other set of questionnaires was to obtain weights of all factors.

Step 2: In-lab driving simulator tests

The in-lab driving simulator is widely used due to its splendid advantages. It is a much safer, ethical, and practical way to study drivers' distractions. The driving simulator used in the research has a half-sized vehicle body, with six-axis movement and 180-degree visual displays.

Step 3: Surveys

Drivers were asked to finish two set of posterior questionnaires which were prepared in step 1 after their driving tasks.

Step 4: Results and analysis

Data collected from both posterior questionnaires and driving test records were analyzed in the statistics tool SPSS. Analysis of Variance (ANOVA) test was used to test the differences among drivers' workload on different sign structures.
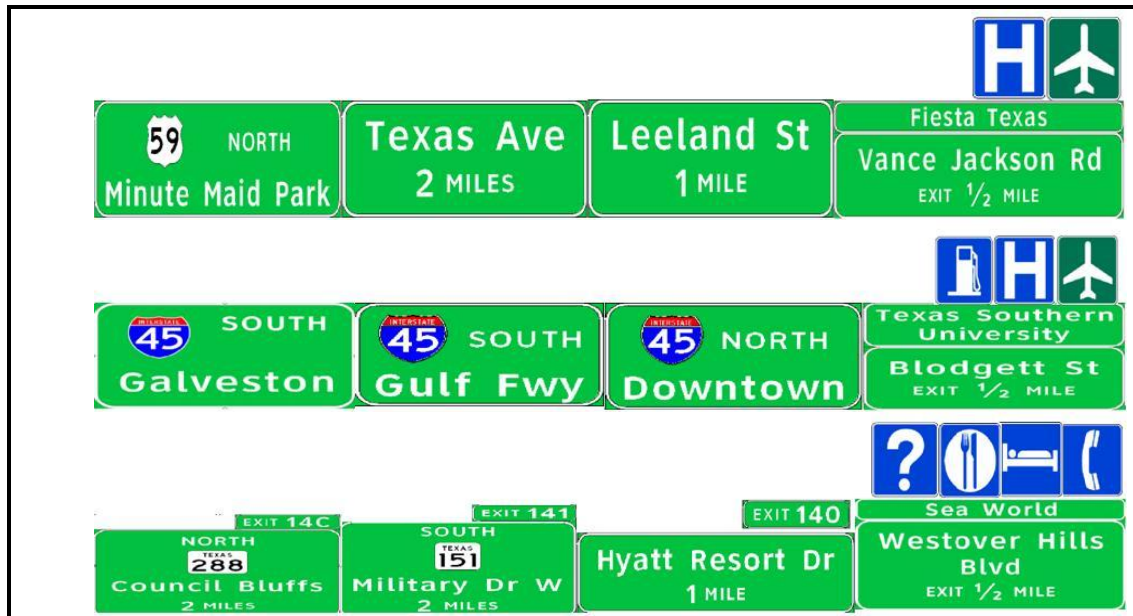
TEST DESIGN

Guide Signs and Scenarios design

      Following MUTCD, 12 guide sign structures were designed for the in-lab simulator tests. The guide sign structures that have the same number of sign panels were ranged as the same sign group. In this way, there are totally 4 groups each has 3 different guide sign structures. From the starting point to the ending point, each scenario has 4 different guide sign structures which are randomly selected from sign groups without repetitiousness. Due to the limitation of the simulator in use, there is no scenario placing 5 sign panels on a single sign structure.

Table 2. List of the Test Sign Panels with Symbol Signs

Group 1 (1 Sign Panels)	Group 2 (2 Sign Panels)

Group 3 (3 Sign Panels)

Group 4 (4 sign panels)

Note: The image of each sign in Table 2 is not exactly the same as what was tested in DriveSafety, where much better quality was designed using standard sign design software SignCAD.

The total information units for each guide sign structure is listed in Table 3, which strictly followed the guideline about maximum amount of information per freeway guide sign structure from MUTCD (19).

Table 3. Sign Structures for In-lab Simulator Tests

Guide Sign Structures	1	2	3	4	5	6	7	8	9	10	11	12
Number of Sign Panels	1	1	1	2	2	2	3	3	3	4	4	4
Number of Symbol Signs	2	3	4	2	3	4	2	3	4	2	3	4
Information Unit Limit	\	\	\	16	16	16	18	18	18	20	20	20
Total Information Units	6	7	8	8	10	12	10	13	18	13	16	20

Questionnaire with Subjective Rating

The questionnaire was designed mostly according to NASA-TLX. There are 6 factors in the rating scale. (1) **Mental Demand** (MD, How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex, exacting or forgiving?) (2) **Physical Demand** (PD, How much physical activity was required?). (3) **Temporal Demand** (TD, How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?) (4) **Performance** (OP, How successful do you think you were in accomplishing the goals of the task set by the experimenter?) (5) **Effort** (EF, How hard did you have to work to accomplish your level of performance?) (6) **Frustration Level** (FR, How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?). One question was

designed for each factor except for Physical Demand, which was obtained from the breaking record of driving tests, and the drivers were asked to rate all other 5 factors after finished each driving task.

Pair-wise Comparisons of Factors

      The six factors mentioned above were defined as Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (OP), Effort (EF), and Frustration (FR). They were 15 combinations shown in Table 4.

Table 4. Comparisons of Six Factors

MD/PD	MD/TD	MD/OP
MD/EF	MD/FR	PD/TD
PD/OP	PD/EF	PD/ER
TD/OP	TD/EF	TD/ER
OP/EF	OP/ER	EF/ER

      Based on this table, a questionnaire was designed for drivers to rate the weight of these 6 factors after finished all tasks. The drivers were required to check the more important one from each pair. The average mean value of each factor is counted as the weight of factors. The average weight by n_d drivers for the fth factor is:

$$W_f = \frac{1}{n_d}\sum_{d=1}^{n_d}W_f^d \text{ , where } W_f^d \in [0, n_f]$$

      The mean weighted workload score of the dth driver to a sign structure with i pieces symbol signs above j pieces sign panel is $MWWL_{ij}^d$, where n_d is the number of drivers, i is the number of symbol signs, j is the number of sign panels. The workload to the fth factor by the dth person to a sign structure with i pieces symbol signs above j pieces sign panel is: L_{ij}^{df} . Then the weighted workload to the fth factor by the dth person to a sign structure with i pieces symbol sign above j pieces sign panel is: W_{ij}^{df} .

Tests Procedure

      In total, there were 24 test drivers who had valid driver licenses whom participated in this test during August 21, 2008 to August 25, 2008. Each test driver was asked to drive through each scenario toward 3 different destinations, and may have been required to reach the same destination twice but in different scenario.

Thus each driver would have 4 tasks in the driving test, which makes it 24 (drivers) * 4 (tasks) = 96 study cases in total for this research. Each time when the driver completed the driving task for one destination, the driver was asked to fill out a related subjective rating questionnaire immediately. Then the driver could continue the next driving task to find the next destination. After all tasks were finished, the drivers were asked to rate the weights of six factors affecting their workload based on their driving experiences.

Test Results and Analysis

A total of 24 participants recruited, with 54.17% female and 45.83% male. Table 5 lists the weights of the six factors applied to drivers' workload assessment. The results were based on the pair-wise comparisons of factors questionnaire after the driving tests from 24 participants' rating. As shown in Table 5, TD (Temporal Demand) has the highest weight (3.67), while Physical Demand (PD) is rated as the lowest one (1.13). That means, most participants deem time pressure brings more workload than other factors. The scores of the Mean Weighted Workload (MWWL) for each guide sign structure is in Table 6, which is calculated based on:

$$MWWL_{ij} = \frac{1}{n_{ij}} \sum_{d=1}^{n_{ij}} MWWL_{ij}^{d} \text{ , where } MWWL_{ij}^{d} = \frac{1}{\sum_{f=1}^{6} W_f} \sum_{f=1}^{n_f} R_{ij}^{df} W_f$$

There is basically a growth trend of workload scores with the increase of sign panels, and with the number of symbol signs. The ANOVA test was used to compare the means of drivers' workload under different scenarios in the research. As this research tries to find out whether the different combination of sign panels and symbol signs will affect drivers' workload and how does it affect drivers' workload, where the number of sign panels and the number of symbol signs are treated as two different independent variables, therefore, the Two-Way ANOVA test was selected that can simultaneously assess the effects of two (or more) independent variables on a single dependent variable within the same analysis as the method of data analysis.

All P_r values are lower than 0.05, there is interaction between "i" and "j" , where "i" represents the number of sign panels and "j" represents the number of symbol signs. There is no significant difference of drivers' workload if there are no more than 2 sign panels for each freeway guide sign structure. When the number of sign panels for each freeway guide sign structure increased to 3 and 4, the Pr value equals to 0.0424 and 0.0060, respectively. That means there is significant difference of drivers' workloads when the sign panels are more than 3.

The ANOVA test also indicates that when the number of symbol signs equals to 2 or 3, there is no significant difference on drivers' workload. While the number of symbol signs become 4, the value of P_r=0.0014, which is much lower than 0.05. This means there is significant difference with drivers' workload in this case. It is concluded that either there are 3 or 4 symbol signs, or there are 4 number of sign panels in one guide sign structure, the workload of drivers would be significantly higher. The worst case is that there are 4 symbol signs and 4 sign panels in one sign structure. In this case, the total number of information unit on a sign structure is 20, which reaches the maximum number of information unit in Table 1.

CONCLUSIONS

This study presents how freeway guide signs would affect drivers' workload. Based on the NASA-TLX, when there are more than 2 guide sign panels or more than 3 symbol signs on the same freeway guide sign structure there is a significant difference on drivers' workload. That means the most desirable numbers of guide sign panels for one freeway guide sign structure are 1 and 2, while 1 to 3 for small symbol signs. This can be used as a reference for guide sign placement on freeways.

REFERENCES

Cai, H., Lin, Y. and Mourant, R. R. (2007). Evaluation of Driver Visual Behavior and Road Signs in Virtual Environment，PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 51st ANNUAL MEETING— 2007.

Clark, J. and Funkhouser, D. (2007). Comparing Data Collection Scenarios used for Examining Drivers' Sign Viewing Behavior: Closed Course versus Open Road Environments.18th Biennial Transportation Research Board Visibility Symposium College Station, TX, April 2007.

Federal Highway Administration (FHA). (2000). *Manual on Uniform Traffic Control Devices*. 2009.

Funke, G.J., Matthews, G., Warm, J.S. and Emo, A. (2007). Vehicle automation: A remedy for driver stress? PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 51st ANNUAL MEETING. Ergonomics.2007.

Hancock, P. A. and Meshkati, N. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload (pp. 139-183). Amsterdam: North Holland Press. 1988.

Hart, S. G. (2007). NASA-TASK LOAD INDEX (NASA-TLX); 20 YEARS LATER, NASA-Ames Research Center, 2007.

Jahn, G., Oehme, A., Krems, J. F., and Gelau, C. (2005). Periheral Detection as a Workload Measure in Driving: Effects of Traffic Complexity And Route Guidance System Use in a Driving Study. Tranportation Research Part F 8, 255-275, 2005.

Lottridge, D. and Chignell, M. (2007). Driving under the influence of phones: the importance of cognitive ability and cognitive style on interruption-related performance, PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 51st ANNUAL MEETING, 2007.

Manual of Traffic Signs and Markings (MTSM), Part 3: Motorway and Expressways. Section 4: supplement destination signs.

Ministry of Transportation and Highways (MTH). (2000). *Manual of Standard Traffic Signs & Pavement Markings*, Table 4.2 Criteria for Supplemental Signing for Major Traffic Generators, 2000.

Ministry of Transportation and Highways of British Columbia (MTHBC). (2000). Canada: Engineering Branch. BC Ministry of Transportation & Highways, *Manual of Standard Traffic Signs & Pavement Markings*, September, 2000.

Minnesota Department of Transportation (MnDOT). (2000). *Minnesota Traffic Engineering Manual*, Chapter 6, Traffic Signs, July 1, 2000.

Missouri Department of Transportation (MoDOT). (2006). Freeway and Expressway Guide Signs - Engineering Policy Guide, 903.8.39 Traffic Generator Signing (MUTCD Section 2E.32a), 2006.

Pline, J. L. (1992). Traffic engineering handbook. Institute of Transportation Engineers (4th ed). Englewood Cliffs, NJ: Prentice-Hall. 1992.

Pratt, S. G. (2003). Work-related roadway crashes: challenges and opportunities for prevention. Technical report: NIOSH Hazard Review, 2003.

Reimer, B., Mehler, B.L., Pohlmeyer, A. E., Coughlin, J. F. and Dusek, J. A. The Use of Heart Rate in a Driving Simulator As an Indicator of Age-related Differences in Driver Workload. Advances in Transportation Studies special issue, 2006.

Qiao, F., Zhang, R., and L. Yu. (2008). Advance Guide Sign Placement Considering Road Users' Workloads and Behaviours. Presented at the 4th International Conference on Traffic & Transport Psychology. Aug. 31-Setp. 4, 2008, Washington, DC, USA.

Texas Department of Transportation (TxDOT). (2006). *Signs and Markings Manual*, Section 9, Traffic Generators, Special Events, and Government Offices," 2006.

Thiffault, P. and Bergeron, J. (2003). Monotony of road environment and driver fatigue: A simulator study. Accident Analysis and Prevention, 35, 381-391. 2003.

Turksen, I. B., Moray, N. and Fuller, K. (1985). A linguistic Rule-Based Expert System for Mental Workload. Department of Industrial Engineering, University of Toronto, 1985.

Wood, C., Torkkola, K. and Kundalkar, S. (2004). Using Driver's Speech to Detect Cognitive Workload. SPECOM, 9TH Conference Speech and Computer, St. Petersburg, Russia, 2004.

Wu, C. G., Tsimhoni, O., and Liu, Y. L. (2007). Development of an Adaptive Workload Management System Using Queueing Network-Model of Human Processor, The 51st Annual Conference of the Human Factors and Ergonomics Society, 2007.