The Effects of Work Zone Configurations on Physiological and Subjective Workload

Mahmoud Shakouri Hassanabadi
Louisiana State University and Agricultural and Mechanical College, mshako1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses
Part of the Mechanical Engineering Commons

Recommended Citation
Shakouri Hassanabadi, Mahmoud, "The Effects of Work Zone Configurations on Physiological and Subjective Workload" (2015). LSU Master's Theses. 3584.
https://digitalcommons.lsu.edu/gradschool_theses/3584

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
THE EFFECTS OF WORK ZONE CONFIGURATIONS ON PHYSIOLOGICAL AND SUBJECTIVE WORKLOAD

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Industrial Engineering in

The Department of Mechanical and Industrial Engineering

by

Mahmoud Shakouri Hassanabadi
B.S., Islamic Azad University, 2008
M.S., Universiti Teknologi Malaysia, 2012
December 2014
I would like to express my special appreciation and thanks to my advisor Dr. Laura Ikuma, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless. I would also like to thank my committee members, professor Dr. Aghazadeh and Dr. Nahmens for serving as my committee members even at hardship. Thank you for letting my defense be an enjoyable moment, and for your constructive comments and suggestions. Finally, I would like to thank Dr. Ishak for letting me use the driving simulator for my research.

A special thanks to my family. Words cannot express how grateful I am to my father, my mother, and my brother for all of the sacrifices that you’ve made on my behalf. Your prayer for me was what sustained me thus far.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... ii

LIST OF TABLES ....................................................................................................................... v

LIST OF FIGURES ................................................................................................................... vi

ABSTRACT ................................................................................................................................. viii

1 INTRODUCTION .................................................................................................................. 1  
   1.1 Background ......................................................................................................................... 1  
   1.2 Problem Statement .............................................................................................................. 3  
   1.3 Objectives .......................................................................................................................... 3  
   1.4 Hypothesis ......................................................................................................................... 4  
   1.5 Significance ......................................................................................................................... 5

2 LITERATURE REVIEW ......................................................................................................... 6  
   2.1 Highway Work Zone .......................................................................................................... 6  
   2.2 Merging Strategies ............................................................................................................ 7  
      2.2.1 Conventional Merge ...................................................................................................... 7  
      2.2.2 Early Merge ............................................................................................................... 8  
      2.2.3 Late Merge ................................................................................................................. 10  
      2.2.4 Zipping ....................................................................................................................... 11  
      2.2.5 Always Close Right Lane ......................................................................................... 13  
      2.2.6 Joint Merge ............................................................................................................... 13  
   2.3 The Concept of Mental Workload ..................................................................................... 16  
   2.4 Driver workload ................................................................................................................. 19  
   2.5 A Model of Mental Workload, Task Performance and Demands ....................................... 21  
   2.6 Measures of Workload ..................................................................................................... 25  
      2.6.1 Performance Measures ............................................................................................... 25  
      2.6.2 Subjective Measures ................................................................................................. 26  
      2.6.3 Physiological Measures ............................................................................................. 27

3 METHODOLOGY .................................................................................................................. 33  
   3.1 Study Design ..................................................................................................................... 33  
   3.2 Participants ......................................................................................................................... 33  
   3.3 Tools .................................................................................................................................. 34  
   3.4 Experimental Model ......................................................................................................... 36  
   3.5 Independent and Dependent Variables ............................................................................. 39  
   3.6 Procedure ........................................................................................................................... 41  
   3.7 Data Analysis ..................................................................................................................... 42  
      3.7.1 Pre-processing .............................................................................................................. 42  
      3.7.2 Statistical pre-processing ........................................................................................... 45  
      3.7.3 Hypothesis Testing ...................................................................................................... 47
LIST OF TABLES

Table 2.1 Factors affecting workload ................................................................. 21
Table 2.2 NASA-TLX Rating Scale and Definitions ........................................... 27
Table 2.3 Summary of HRV Time domain parameters (Tarvainen et al., 2014) ......... 30
Table 2.4 Summary of HRV frequency domain parameters (Tarvainen et al., 2014) .. 31
Table 2.5 Summary of nonlinear HRV methods (Tarvainen et al., 2014) ................. 32
Table 3.1 Physiological measures of workload ..................................................... 39
Table 3.2 Subjective workload variables from the NASA-TLX ............................ 40
Table 3.3 Performance variables ........................................................................ 41
Table 3.4 list of transformed variables ................................................................ 47
Table 4.1 Summary statistics (means and standard deviations) for physiological measures of workload at rest and in three driving scenarios (N=60*) .............................................................. 50
Table 4.2 Summary statistics (means and standard deviations) for physiological measures of workload at rest and in different traffic densities (N = 90+) ............................................................ 52
Table 4.3 Summary statistics (means and standard deviations) for NASA TLX components in three driving scenarios (N=60) ........................................................................ 54
Table 4.4 Comparison of NASA TLX components (means and standard deviations) in two traffic density levels (N=90) ............................................................... 55
Table 4.5 Descriptive statistics (means and standard deviations) for performance variables in three driving scenarios (N=60) ................................................................. 57
Table 4.6 Comparison of performance variables (means and standard deviations) in two traffic density levels (N=90) ........................................................................ 58
Table 4.7 Pearson correlation matrix for physiological measures of workload, subjective workload and performance variables ...................................................... 61
LIST OF FIGURES

Figure 2.1 Conventional merge design layout (U.S. Department of Transportation, 2009) ........................ 8
Figure 2.2 Dynamic early merge design layout (McCoy & Pesti, 2001) ................................................. 9
Figure 2.3 Late merge design layout (Pesti et al., 1999) ............................................................................. 11
Figure 2.4 Zipper sign (Risten) in the Netherlands ..................................................................................... 12
Figure 2.5 (a) MUTCD W4-2 (b) Experimental merge sign ................................................................. 12
Figure 2.6 Joint lane merge configuration layout ....................................................................................... 14
Figure 2.7 Graphical representation of the Yerkes-Dodson Law ........................................................ 22
Figure 2.8 Task demand and performance relationship (Meister, 1976) .................................................. 23
Figure 2.9 Relationship between workload and performance: six theoretical levels (de Waard, 1996) ........................................................................................................................................ 24
Figure 2.10 Result of G*power estimates for the number of required participants in this study ... 34
Figure 3.1 View of driving simulator ........................................................................................................ 35
Figure 3.2 Polar RS 800 heart monitoring watch ....................................................................................... 36
Figure 3.3 Developing construction zone layout for driving simulator ...................................................... 36
Figure 3.4 Conventional merge layout with right lane closure ............................................................. 38
Figure 3.5 Joint lane merge configuration ................................................................................................. 38
Figure 3.6 Sample exported RR data from PolarPro 5 ........................................................................ 43
Figure 3.7 Example Kubios HRV interface ............................................................................................... 44
Figure 3.8 Example Kubios HRV time domain measures ........................................................................... 44
Figure 3.9 Example Kubios HRV frequency domain measures ............................................................. 45
Figure 4.1 Mean physiological measures of workload in different driving scenarios (a) RMSSD; (b) LF; (c) HF; (d) LF/HF .................................................................................................................................. 51
Figure 4.2 Mean physiological measures of workload in different traffic densities (a) RMSSD; (b) LF ....................................................................................................................................... 52
Figure 4.3 Mean percent subjective workload for different driving scenarios .......................... 54
Figure 4.4 Mean percent subjective workload in different traffic densities ....................... 56
Figure 4.5 Mean variability for performance variables for different driving scenarios ........... 57
Figure 4.6 Mean variability for performance variables in different traffic densities .............. 58
Figure 4.7 Driving scenario-traffic density interaction graph for brake variability ................. 59
Figure 4.8 Driving scenario-traffic density interaction graph for speed variability ............... 59
The objectives of this research were to (a) study the effect of work zones and traffic density on physiological measures of workload, subjective workload and performance variables (b) study the relationship between physiological measures of workload, subjective workload and performance variables. Conventional lane merge (CLM), joint lane merge (JLM) and a road without a work zone (control) were modeled with high and low traffic density by using a full-size driving simulator. 13 female and 17 male students volunteered to participate in this study. Data regarding physiological measures of workload through heart rate variability measures (RMSSD, LF, HF and LF/HF ratio) were collected by using a heart monitoring watch. NASA-TLX was used to measure subjective workload. Variability in steering, braking and speed were used as performance variables. Results showed that the driving scenarios and traffic density did not affect physiological measures of workload. In terms of subjective workload, CLM and JLM did not differ significantly from each other. However, with respect to mental demand, temporal demand, effort and total workload, CLM was significantly more demanding than the control group. Total workload for driving in high traffic density was 27.2% more than that of in low traffic density. No significant differences were observed in brake variability between different scenarios. However, CLM and JLM had significantly higher speed variability than the control group but they were not significantly different from each other. Steer variability and brake variability were higher in high traffic density. In conclusion, physiological measures of workload showed no sensitivity to changes in the work zone but subjective and performance variables are influenced and can be used to compare different work zone configurations.
1 INTRODUCTION

1.1 Background

It is not always feasible to stop traffic in order to perform maintenance on a road. The common practice is to close only the lane where maintenance is going on and guide the moving vehicles to the open lane during the maintenance period. This period can vary from several hours to several weeks and is dependent on the type of work and conditions of the site. One of the challenges faced by transportation officials and road contractors regarding work zones is reducing the negative impacts of work zones on traffic flow and providing a safe environment for workers and motorists. Inefficient planning for traffic operation control near work zone areas can lead to high traffic queues, increased number of forced merges and increased chances of roadway accidents (Al-Kaisy & Hall, 2002). According to the National Center for Statistics and Analysis (2010), in 2010, 87,606 accidents happened in active work zones with lane closure in which on average 46.9% of them were rear-end collisions, 20.3% of them were fixed object collisions, 13.6% of them were sideswipe and the rest were other type of accidents that happened during the day. Of the 87,606 accidents, 720 were fatal. 22% of fatal work zone crashes occurred on urban interstates and 59% of fatal work zone crashes occurred on roads with speed limit of 55 mph or more. On average, 85% of deaths in work zones were drivers and passengers in cars (FHWA, 2013).

With the increase in the number of cars and highway networks, there is a growing concern for road safety. In recent years, several researchers have studied the efficiency of different merge configurations near work zones in terms of metrics such as throughput, number of forced merges (Rouphail et al., 1988), vehicles operating speed (Beacher et al., 2005), deceleration (Ishak et al., 2012), travel time (Rayaprolu et al., 2013), and other traffic flow...
characteristics (Grillo et al., 2008; Jiang, 2007; McCoy & Pesti, 2001; Morgan et al., 2010; Pesti et al., 1999). Previous findings show that most road traffic accidents can be partially attributed to human factors (Öz et al., 2013). However, few researchers have studied the psychological and physiological impacts of such configurations on drivers. Driving requires performing physical and cognitive tasks under time pressure which makes driving through work zones physically, mentally, and temporally demanding. Drivers’ mental workload, for instance, is one of the factors that may influence drivers’ behavior in highway work zones (Brookhuis & de Waard, 2010). Workload can be simply defined as the demand placed upon humans while performing a task. Inadequate workload (either too low or too high) may lead to insufficient attention, wrong perception and inadequate information processing (Brookhuis & de Waard, 2010; Harrison & Fillmore, 2011; Leung & Starmer, 2005). As workload increases, performance in information processing degrades (Gawron, 2008). Research shows that high workload leads to error, impaired performance and loss of situation awareness; and low workload negatively affects vigilance and alertness (Hockey, 1997; Verwey & Zaidel, 1999). High road-environment demands (e.g., having to merge in heavy traffic) increase workload, while the effects of alcohol, persisting monotony and fatigue increase workload by a reduction in capacity (de Waard, 1996; Schneider et al., 1984; Wierwille & Eggemeier, 1993). In order to ensure safety, comfort, and long-term efficiency of drivers in work zones, task demands need to be regulated so that drivers can perform merging maneuvers efficiently without being overloaded. Hence, understanding how merge configurations affect drivers’ workload and response to changes in the driving environment is a crucial step in improving work zone safety. Measuring mental workload while driving may provide an indication of the cognitive demands placed on the driver (Brookhuis & de Waard, 2010; Öz et al., 2013).
Few of the published works in the transportation and safety domain have studied the effects of a new merge configuration on both physical and psychological aspects of driving. Research on drivers’ workload in different environments can provide an understanding of the underlying mechanisms that result in accidents whereas, studying the factors affecting performance variables helps to improve traffic flow. This paper investigates the effects of work zone configurations and traffic density on performance variables, physiological measures of workload, and subjective workload. The main objectives of this research are to (a) study the effect of work zones and traffic density on physiological measures of workload, subjective workload and performance variables (b) study the relationship between physiological measures of workload, subjective workload and performance variables.

1.2 Problem Statement

Literature regarding lane merge configuration is replete with studies focusing on the operational aspects of merge configurations such as operating speed, throughput, delays, etc. However, few researchers have studied the psychological and performance impacts of such configurations on drivers. This paper is an attempt to fill this gap by studying the effect of work zone configurations on physiological and subjective measures of workload.

1.3 Objectives

The main objective of this research is to investigate the effects of work zone configurations on the physiological and subjective measures of workload and whether there is any association between subjective workload and physiological measures of workload. It specifically aims at answering the following research questions:

1) How do different work zones affect subjective and physiological measures of workload?
2) Is there any association between subjective and physiological measures of workload?
3) Is there any association between workload and performance variables?

To answer these research questions the following steps should be undertaken:

1) Design work zones and collect data using a driving simulator with regard to physiological
   and subjective measures of workload along with information related driving behavior
   (such as speed, lane change pattern, braking pattern, etc.).
2) Quantify the impacts of different work zone configurations on the physiological and
   subjective measures of workload and performance variables.
3) Examine the associations between drivers’ physiological workload, subjective workload
   and performance variables.

1.4 Hypothesis

This study has three hypotheses:

- Hypothesis 1:
  - $H_0$: There is no difference in drivers’ physiological workload at the presence of
    work zones in different traffic densities.
  - $H_a$: There is at least a difference in drivers’ physiological workload at the
    presence of work zones in different traffic densities.

- Hypothesis 2:
  - $H_0$: There is no difference in drivers’ subjective workload at the presence of work
    zones in different traffic densities.
  - $H_a$: There is a difference in at least one measure of drivers’ subjective workload at
    the presence of work zones in different traffic densities.
• Hypothesis 3:
  
  o $H_0$: There is no association between subjective workload, physiological workload and performance variables
  
  o $H_a$: There is at least an association between subjective workload, physiological workload and performance variables

1.5 Significance

The current research investigates driver safety which is a critical component of transportation systems. This research will develop a better understanding of how work zone configurations affect workload which can influence driver safety. The novelty of this study lies in the usage of workload as a measure of safety to compare different work zone configurations.
2 LITERATURE REVIEW

2.1 Highway Work Zone

A highway work zone is a part of the road where construction or road maintenance takes place. Work zones impede traffic flow and create congestion. In order to keep the continuity of movement for motor vehicles, temporary traffic control plans (TTC) should be used. Some of these plans are introduced in the Manual of Uniform Traffic Control Devices (MUTCD) which is a national standard in the U.S. for traffic control devices used on all public streets and highways (U.S. Department of Transportation, 2009). According to MUTCD, a common TTC includes flaggers, traffic signs, arrow panels and portable changeable message signs, channelizing devices, pavement markings, lighting devices, and temporary traffic control signals (U.S. Department of Transportation, 2009).

Lanes in a typical work zone can be classified into two types: merge lanes and through lanes. A merge lane is the lane that is closed due to road work and a through lane is the one that is left open for vehicles to pass by. Vehicles in the merge lane are expected to complete their merge and go to the through lane before they enter the work zone area. However, studies show that the majority of drivers remain in the merge lane and perform their merging maneuvers in the work zone area which results in traffic congestion and in some cases accidents (Yi & Mulinazzi, 2007). According to a field study of driver behavior near work zones (Steele & Vavrik, 2009) 94.4% of drivers in the merge lane started to change lanes at about 500 feet before the road taper. In general, the merge lane should be long enough so that at least 85% of drivers can complete their merging maneuvers (Ahammed et al., 2008; Makigami et al., 1988). Makigami et al. (1988) developed an analytical method to determine the necessary merging length and concluded that 700 m is an optimal length for a transition section in three and four lane highways.
Inefficient planning for traffic operation control near work zones can lead to high traffic queues, additional fuel consumption, increased number of forced merges and increased chances of roadway accidents (Al-Kaisy & Hall, 2002). Research on improving the operational efficiency of work zones in recent years has led to the advent of new merge configurations. In addition to Conventional Lane Merge (CLM) which is recommended by the United States Department of Transportation (U.S. Department of Transportation, 2009), there are other configurations such as early merge, late merge and zipping that are used in different parts of the U.S. However, despite all the efforts to modify merge configurations and improve work zone safety, the high rate of crashes and fatalities in work zone areas are still unacceptable and the need to examine new merge configurations and improve efficiency and safety of merging maneuvers still exists. New configurations can be designed by using special geometric configurations and advanced signage that lead to improvements in the merging experience of drivers at work zones (Rayaprolu et al., 2013).

2.2 Merging Strategies

This section provides an overview of several studies that evaluated the operational efficiency of the CLM strategy, along with some unconventional lane merge configurations such as static early merge, static late merge, dynamic early merge, dynamic late merge and zipping.

2.2.1 Conventional Merge

The current lane closure design (CLM) specified in the MUTCD (U.S. Department of Transportation, 2009) is the most commonly used design in the U.S. and seeks to guide drivers from the closed lane to the open lane safely. Under the CLM configuration, when two lanes merge into one lane, vehicles in the open lane are given the right of way, while those in the closed lane are expected to move into the open lane before the two lanes merge (Figure 2.1).
Vehicles in the open lane are given the opportunity to continue to move into the work zone area without stopping, but vehicles in the closed lane may have to slow down or stop if the merging gaps in the open lane are limited (Rayaprolu, 2010). However, the safety of this merging configuration is only effective in low to moderate traffic densities (Ishak et al., 2012). Some advantages of the CLM in the U.S. are its widespread usage and drivers’ familiarity with the incorporated traffic signs. However, increased potential for rear end and side swipe crashes and longer queue lengths in high traffic density are the drawbacks of this merge (Ishak et al., 2012).

Figure 2.1 Conventional merge design layout (U.S. Department of Transportation, 2009)

2.2.2 Early Merge

Early merge aims at providing enough response time for drivers approaching a merge by means of placing warning signs in advance of the taper (McCoy & Pesti, 2001). Early merge is divided into static early merge and dynamic early merge. In static early merge, drivers are informed about the upcoming lane closure by advance “LANE CLOSED” signs placed nearly 1.5 miles before the taper. Also, lane reduction signs are placed 1500 ft. before the taper, followed by flashing arrow panels at the beginning of the taper. This type of lane merge is
suitable when demand is below capacity but fails as congestion develops due to speed variation between lanes as drivers in the closing lane tend to pass those in the open lane. Contrary to static early merge where sign distance intervals are fixed, the signs in dynamic early merge are responsive to real time traffic measurements (Figure 2.2). When stopped vehicles are detected by sonic detectors near the signs, a signal is transmitted to the nearest upstream sign. Signs in dynamic early merge are placed at either .25-.5 mile intervals upstream of the lane closure. When the signal is received by a sign, it alerts the drivers by showing a “DO NOT PASS” message. Another difference between early static and dynamic merge is the incorporation of beacon lights in dynamic merge. The lights are deactivated once a stopped queue is no longer detected.

![Figure 2.2 Dynamic early merge design layout (McCoy & Pesti, 2001)](image)

Early merge strategies may be successful in reducing the number of forced merges in the transition area, however, travel times during high traffic density may increase (Rouphail & Tiawari, 1985). Tarko et al. (1998) found that using early dynamic merge strategies increased the
size of queues and length of merging zones due to the reduction of speed in the open lane, especially during high traffic. McCoy and Pesti (2001) found a smooth merging behavior in low traffic with the dynamic early merge, but abrupt decelerations and large queue lengths during high traffic led to a reduction in throughput. Early merge strategies potentially can reduce traffic volume. However, as with the CLM, its efficiency declines in high traffic density, and chances of accidents and aggressive driving increase.

### 2.2.3 Late Merge

The late merge strategy was proposed to reduce aggressive driving behavior between motorists in the closed and open lanes (McCoy & Pesti, 2001) (Figure 2.3). In this strategy vehicles are encouraged to stay in their lanes until they reach the merge section. As like the early merge strategy, late merge is also divided into static late merge and dynamic late merge. The concept behind the late merge is to encourage drivers to use both lanes until a specified merging point. Once vehicles reach the merging point, those in the closed lane merge with vehicles in the open lane in an alternating pattern. Typically, a “Use Both Lanes to Merge Point” sign is placed approximately 1.5 miles (2.4 km) in advance of the taper.

Several researchers studied the efficacy of late merge configuration in terms of traffic flow characteristics and safety in work zones. Beacher et al. (2005) compared the CLM and static late merge configurations and found that except for positive response from drivers towards static late merge, no significant difference in throughput compared to the CLM was found. Similarly, Kang et al. (2006) concluded that the behavior of the dynamic late merge strategy is analogous to the CLM in unsaturated traffic densities. According to McCoy and Pesti (2001) forced merges in the late merge strategy was 75% lower than CLM at high densities. Forced merges occur when there is not enough space between vehicles in the closed lane and open lane.
and as a result, the vehicles in the closed lane attempt to merge with evasive maneuvers. The result also showed 30% fewer lane straddles at densities below 25 vehicles per mile. Finally, a study by Grillo et al. (2008) found that the dynamic late merge configuration is more effective on highways with moderate to heavy congestion prior to construction work zones. As a result, benefits of the late merge lie in its application in high volume traffic. It reduces rear end crashes and creates shorter queues. However, compliance of drivers to this new strategy is low which creates hazards in low volume traffic (Beacher et al., 2005).

![Late merge design layout](image)

Figure 2.3 Late merge design layout (Pesti et al., 1999)

### 2.2.4 Zipping

An alternate merging strategy called “zipping signs” is used in the Netherlands, Belgium and Germany (Figure 2.4). In this strategy, during congested periods, vehicles in the open lane permit adjacent vehicles to merge in an alternating pattern until the congested period ends. Dijker and Bovy (1998) studied the performance of zipping strategy in the Netherlands, and found that compared to other configurations, zipping maneuvers do not affect throughputs in the zipping strategy. In the United States, the Connecticut Department of Transportation proposed a test sign similar to the zipping sign (Feldblum et al., 2005). This sign was the result of two surveys that showed it was the statistically best understood sign among 6 proposed signs.
(Figure 2.5). This test sign was used in the field along with the W 4-2 sign and the results showed that the test sign had statistically increased the desirable number of merges from 56% to 66% and reduced the undesirable merges from 9% to 5%. One advantage of this merging strategy is that speed is better maintained as motorists travel through the merging area (Idewu, 2006).

Figure 2.4 Zipper sign (Risten) in the Netherlands

Figure 2.5 (a) MUTCD W4-2 (b) Experimental merge sign
2.2.5 Always Close Right Lane

This strategy, which is commonly used in Arkansas, advocates for closing the right lane at all times. Drivers who are familiar with the rules know ahead of time which lane is ending. Once the first merge is completed, drivers are channeled to the appropriate side of construction. Although the effects of this type of strategy are not well documented, one study showed that the crash rate in always close right lane configuration was 46% lower than the CLM (Schrock & McClure, 2009). This configuration creates less confusion on which lane is closed and may reduce the number of sideswipe crashes. It is widely recognized that when congestion develops and queues form at the approach to work zones, the risk of crashes increases, especially on major highways where speeds are high and drivers are accustomed to unencumbered travel. Additionally, the problem can be compounded by limited sight distance and roadway curvature. As a result, in high traffic density, increased back-of-queue crash at lane closures in always close right lane strategy presents a very serious safety condition.

2.2.6 Joint Merge

The crash analysis results of work zone areas show that the rate of crashes in advance warning areas where drivers usually perform their merging maneuvers is higher compared to other parts of the road (Bureau of Transportation Statistics, 2011). Therefore, the Joint Lane Merge (JLM) configuration was proposed as an alternative to the CLM configuration (Idewu & Wolshon, 2010) with more emphasis on the configuration of the transition area. In the JLM configuration (Figure 2.6), motorists in both lanes have equal right of way, as opposed to CLM where only the open lane has the right of way. The JLM configuration is divided into five distinct zones as shown in figure 2.6. The advance warning zone in the JLM is typically a mile long and compared to the CLM includes more traffic signs to inform drivers about the upcoming road conditions. At the end of the advance warning zone, two blinking arrow signs are placed on
both sides of the road, suggesting that vehicles should merge by taking alternating turns over the transition area. The transition zone is divided into three sections. In the first section, both lanes are tapered from the full lane width (typically 12 ft) to nearly 6 ft to form a single lane of 12 ft. In the second section, vehicles merge to the center line, and in the third section vehicles are guided by the flashing arrow sign either to the right or left lane, depending on the open lane in the work zone area. The activity and termination areas in the JLM configuration are identical to those in the CLM configuration.

![Figure 2.6 Joint lane merge configuration layout](image)

Figure 2.6 Joint lane merge configuration layout

Several studies evaluated the operational efficiency of joint merge. Idewu and Wolshon (2010) conducted a field study to evaluate the effects of the JLM on traffic in a controlled work zone in Louisiana. The comparison of merging speed between the JLM and CLM showed no significant difference at volumes ranging from 600 to 1,200 vehicles per hour (vph). However, the experimental results did suggest that drivers going through the JLM were more cautious in their merging maneuvers. Ishak et al. (2012) examined and compared the safety performance of the conventional lane merge configuration with joint merge in terms of uncomfortable
decelerations and speed variance by using a microscopic simulation model (VISSIM). Results showed that in most simulation scenarios, for the advance warning zone, the CLM configuration exhibited lower frequency of uncomfortable decelerations as opposed to the JLM configuration. However, for low flow rate of 500 vph, no significant differences were detected. For the transition area, in most scenarios with low to moderate flow rates (500–1500 vph) the JLM configuration had less frequent rate of uncomfortable decelerations and therefore was considered safer than the CLM configuration. In another study, Rayaprolu et al. (2013) compared performance measures in terms of total throughput and average delay time between CLM and JLM. Their results showed that at low levels of demand (500 and 1000 vph) both configurations had similar operational performance in terms of throughput and average delay time. At high levels of demand the JLM had significantly higher throughput and shorter delays than the CLM.

Open literature regarding lane merge configuration is replete with studies focusing on the operational aspects of merge configurations like operating speed, throughput, delays, etc. Despite efforts to modify merge configurations and improve work zone safety, the high rate of crashes and fatalities in work zone areas are still unacceptable which indicates that the current safety measures and applied policies are deficient in reducing risky driving behavior (Hirsch, 2003; Mayhew, 2007).

From a human factors perspective, driving requires performing physical and cognitive tasks under time pressure, and this makes driving through work zones physically, mentally, and temporally more demanding. High demand tasks result in so-called workload overload that may create stress for drivers and increase the risk of accidents (Wickens & Hollands, 2000). In order to ensure safety, health, comfort, and long-term efficiency of drivers in work zones, designers should regulate task demands so that drivers can perform merging maneuvers efficiently without
being mentally, physically and temporally overloaded. However, there is a dearth of information on how drivers react to different work zone configurations. Understanding how drivers respond to changes in the driving environment and what road characteristics trigger risky driving behavior near work zones is a crucial step towards improving work zone safety. The existing literature clearly suggests that many factors determine the efficiency of a merge configuration. These factors are divided into two broad categories: geometric configurations factors and human behavioral factors. The aim of the present study is to determine the effects of merge configuration and traffic density on workload and performance factors such as speed, brakes etc. This research demonstrates the use of human factors analysis techniques to understand of drivers’ behavior and performance in work zones.

2.3 The Concept of Mental Workload

Workload can be simply defined as the demand placed upon humans. However, this definition attributes workload to only external resources. Some scientists prefer the term ‘experienced load’ which connotes task and person-specificity (Rouse et al., 1993). Factors such as motivation to perform a task, applied strategies, individual capabilities and the mood and state of the operator can all play significant roles in affecting experienced load whereas demand is only goal driven and independent of individuals.

Workload studies focus on individual limitations that affect performance (e.g. in terms of accuracy and speed) and methods of attaining task goals (e.g. order of actions). The objective of workload studies is to show how different individuals respond to a specific task and investigate the interaction between operator and task structure (de Waard, 1996). Workload literature is replete with terms that might be used interchangeably in daily life but have different connotation in research studies. Task complexity, task demand and task difficulty are some of these terms
and understanding the differences between these terms is an essential step towards fully appreciating workload studies.

The online Psychology dictionary defines task demands as how hard and how long people need to work to complete a task. As mentioned earlier, task demands are goal-specific and vary according to the objectives of the performed task (Psychology Dictionary, 2014b). Task complexity is the degree of complicated actions needed to complete a task (Psychology Dictionary, 2014a). Complexity increases with an increase in the number of stages of processing that are required to perform a task. Task demand and complexity are mainly external, but both depend upon (subjective) goals set for task performance (de Waard, 1996). Difficulty of a task is related to the amount of resources that an individual uses to perform a task. Task difficulty depends on many factors such as experience, context, priority of tasks and resource allocation and the strategy used to accomplish a task (de Waard, 1996). In this text demand will henceforth be used to indicate the task demands and load or workload will be used to describe the effect the demand on the operator in terms of stages that are used in information processing and their energetics. More specifically, workload is the specification of the amount of information processing capacity that is used for task performance.

Kantowitz (1987) has defined complexity as a property of a task in isolation and difficulty as a property of a task in interaction between individual and task. As an example in a math exam, the passing score is the same for everyone and depends on the number of correctly answered questions. However, goal setting (e.g. getting A or B in the exam) determines the amount of task demand. On the other hand, the difficulty of calculations depends on the individual who is taking the exam (e.g. how much he/she is prepared, whether he/she had enough
sleep the night before the exam, etc.). Furthermore, solving the questions for someone who has practiced a lot and is more experienced with math problems is much easier than a novice person.

Thus, the goal of workload measurement is to specify how much capacity is used in performing a task. In this definition, workload is not solely dependent on the task but also on the amount of available resources or the amount that the operator is willing to allocate to do the task (Meijman & O’Hanlon, 1987; Zijlstra & Mulder, 1989). Mobilization of additional resources as a compensatory process is called effort which is of a primary importance (Aasman et al., 1987; Vicente et al., 1987). This is due to the fact that firstly, the expended effort on a task is not necessarily related to the task demand and an operator can choose how to react to the demand based on his/her goals or other criteria (Vicente et al., 1987). Secondly, performance and amount of invested effort are not always related to each other. Task structure is an important factor which affects the amount of expended effort (Norman & Bobrow, 1975). Furthermore, experience and operator’s familiarity with the task can influence the amount of expended effort. According to (Mulder, 1980) mental workload is linked to the controlled mode of information processing. Schneider and Shiffrin (1977) proposed two modes of information processing: automatic versus controlled information processing. Automatic processing, as opposed to controlled, is fast and not conscious and requires low amount of resources. Controlled processing, on the other hand, requires effort, conscious but flexible. In controlled processing, information should be retained in the working memory for further analysis. Hence, it requires resources and attention. Mulder (1980) defines the amount of time that it takes for a person to process information through controlled mode as mental effort. According to (Meijman and O’Hanlon (1984)) there is direct relationship between mental demand and controlled processing...
time. As mental demand increases, so does processing time. Workload assessment is coupled with task difficulty experienced by an operator (Gopher & Donchin, 1986)

Operators can adapt their behavior and cope with an increase in demand. They can also change their strategy and task goals and accept a lower performance level or they can give up completely (de Waard, 1996). Strategies will differ between individuals, and some strategies will be more effective and require less effort to reach the same level of performance. In other conditions in which a change in strategy or ‘quitting’ behavior occurs, measures of effort may remain unchanged or even show a decrease, while performance measures will indicate decreased task performance.

### 2.4 Driver workload

One of the models that researchers use to evaluate workload is measuring performance in doing a main task. Based on this model, as workload increases, additional information processing degrades (Gawron, 2008). In the driving context, primary task can be defined as safe control of vehicle in traffic (Parkes, 1991). Driving is a dynamic task and is influenced not only by the state of the driver but also by external factors such as weather, road conditions and ambient traffic behavior. Hence, modeling driving behavior is not an easy job. Michon (1985) proposed a model for driving tasks which involved at least three levels. In the first level, which is called strategic level, strategic decisions such as route choice, stop points etc. are made. At the intermediate level, strategies such as maneuvering level and type of reactions to external factors are decided. And finally, at the bottom level, which is known as control level, the basic vehicle controls such as lateral control, headway distance control, etc. are determined. The lowest level as opposed to other two levels which require higher level of controlled processing is more automatic. The driver-performance model can be used in any of these three levels to measure workload. For
example, steering-wheel movements reflect performance at the lowest level, car following performance and mirror looking are processes at the maneuvering level, while errors in route choice reflect performance at the strategic level (de Waard, 1996).

Any increase in the demand at any level can affect performance in that level and subsequent levels. As an example, a novice driver cannot perform all control tasks at the same time and as a result may experience higher workload compared to an experienced driver. This may result in performance deficiency in fulfilling the tasks in other levels (e.g. missing a stop sign, reduction in speed etc.).

Sources of driver workload may be found both inside and outside the vehicle. Driving through a junction in rush hour or using a navigation system while conversing with someone in the car can be examples of external and internal vehicle workload sources. Driving heavily relies on the use of visual modality which means a driver to a large extent performs visual activities (e.g. scanning for obstacles, hazards, road signs, etc.). Therefore, any activities that use the same modality, in this case visual, consume the available resources and as a result, degradation in performance occurs. For example, looking at the navigation system while driving interferes with the driving task more than speaking with a passenger in the vehicle. In the former one, both driving and looking for a correct route on the navigating system require the use of visual resources whereas in the latter one, conversing with a passenger uses a different modality and interference is minimal. In recent years, a plethora of research has been conducted on the effect of in car technologies on driver workload (Consiglio et al., 2003; Engström et al., 2005; Jahn et al., 2005; Liu & Lee, 2005, 2006). Results show that these devices have negative impacts on workload and expose drivers to risky situations. Although the use of in car technologies have
their own merits, such devices result in information processing overload, which can lead to increased mental workload (Verwey, 1990).

Table 2.1 displays factors that influence workload. These factors may either increase or decrease workload. In general, feedback is intended to reduce demand, but sometimes it increases workload by providing additional information that has to be processed. High road-environment demands, e.g., having to merge in heavy traffic increases workload, while the effects of alcohol, persisting monotony and fatigue increase workload by a reduction in capacity (de Waard, 1996; Schneider et al., 1984; Wierwille & Eggemeier, 1993).

<table>
<thead>
<tr>
<th>Table 2.1 Factors affecting workload</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Driver State Factors</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Driver Trait Factors</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Environmental Factors</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

2.5 A Model of Mental Workload, Task Performance and Demands

Task performance is an objective measure that specifies how well an individual is performing. Task performance differs from task difficulty which is the experience that an individual goes through when performing the task. Many factors such as task complexity, the operator state and capabilities can influence task difficulty. As the task becomes more difficult, more resources are allocated, and mental workload measures the amount of these allocated
resources. One of the earliest models developed to show the relationship between performance and workload is known as Yerkes-Dodson law (Yerkes & Dodson, 1908). The Yerkes-Dodson law describes the relationship between arousal and performance as shown in Figure 2.7. The Yerkes-Dodson law depicts a drop in operator performance when the arousal in the task is low or high.

![Graphical representation of the Yerkes-Dodson Law](image)

Figure 2.7 Graphical representation of the Yerkes-Dodson Law

Figure 2.8 shows that performance in Meister’s model is stable in regions A and independent of task demand in region C. However, it is only sensitive to changes in demand in region B. As a drawback to Meister’s model, it is difficult to identify where the region of underload occurs. This can be solved by adding a deactivation region (region D) before region A. This region shows the effect of monotonous, low demand tasks that can consume resource capacity or impede the allocation of resources and add up to task difficulty and workload consequently (de Waard, 1996; Meijman & O’Hanlon, 1984; O’Hanlon, 1981).
Another important issue regarding Meister’s model is the determination of workload redline or in other words how much workload is too much? (Reid & Colle, 1988; Wierwille & Eggemeier, 1993). Rueb et al. (1992) define this line as the transition from region A to region B. This is the point where performance starts degrading or personal well-being is being affected. Reid and Colle (1988) used self-report ratings of workload to measure performance decrements to identify workload redline. Their result was similar to those of Rueb et al. (1992) in which performance decrements occur at the transition from region A to B.

Meister’s model of task demand and performance consists of three regions. However, de Waard (1996) proposed an improved version of this model which consists of six regions. This model is depicted in figure 2.8. In region D, although the task demand is very small but due to the operator’s condition being affected the performance is poor. For example the operator in region D might be tired or distracted and therefore unable to cope with tasks with minimal

---

Figure 2.8 Task demand and performance relationship (Meister, 1976)
demands. Performance is high in regions A1 to A3 and reaches its optimal state, with more effort, in region A2. The operator in these regions have enough capacity and resources to cope with task demand and as a result the workload in these regions is low. Task demand in region B starts to exceed capacity and therefore on the one hand, the operator’s performance starts to decline and on the other hand, workload starts to increase in this region. Performance reaches its lowest level in region C. In this region the operator is overloaded and in order to restore performance, the task demand should be reduced.

Figure 2.9 Relationship between workload and performance: six theoretical levels (de Waard, 1996)

In order to guarantee driver’s safety in work zones, it is essential to keep a driver in high performance regions (region A in Meister’s model and regions A1 to A3 in de Waard’s model). As a result, measuring how work load is affected in different work zone configurations and
identifying a proper configuration that promotes high performance is an important step to mitigate work zone risk factors.

2.6 Measures of Workload

Three types of workload measurements have been widely used in the literature to evaluate the amount of workload imposed on an operator by the designed task. These measures can be categorized as performance measures (objective measures), subjective ratings and physiological parameters.

2.6.1 Performance Measures

Workload can be measured by studying how performance is affected if a criterion in the task is changed. One drawback to performance measures is the difficulty in predicting workload in regions D and C in de Waard’s model. In these regions high effort is usually perceived as high workload and low performance. Therefore, it is better if performance measures are accompanied by other types of measures for evaluating workload.

Performance measures can be divided into three groups; primary task- measures, secondary-task measures and reference tasks. Primary-task measures are usually used in either a laboratory or field setting. In primary-task studies, a subject’s performance is measured based on specific performance criteria such as number of errors, speed, reaction time, etc. (Baldauf et al., 2009; Brookhuis et al., 1985; Cantin et al., 2009; Green et al., 1993).

In secondary-task measures an additional task is introduced and workload is measured while the operator is responding to multiple tasks (Ma & Kaber, 2005; Teh et al., 2014). In this case it is assumed that due to the introduction of the new task, the operator consumes the spare capacity and as the result he/she should exerts more effort to compensate for the resources. As a result, it is deemed that secondary task can increase workload which leads to performance
degradation. One criticism to this method is the interference of secondary task with the primary task. Moreover, the secondary task must be demanding enough to ensure that performance on it is indicative of spare capacity (Bortollussi et al., 1987; Jahn et al., 2005). Reference tasks are standardized laboratory tasks measured before and after the task under evaluation (Benedetto et al., 2011; Birrell & Young, 2011). They mainly serve as a technique for assessing trends in primary task performance. The change of performance on the reference task over time indicates the effects of the mental load produced by the primary task. If subjective and physiological measures are added to the reference task, the effort needed to maintain performance on the primary task could also be inferred, particularly when the operator’s state is affected. The use of standard reference task batteries is common in organizational psychology (c.f., Van Ouwerkerk et al., 1994).

2.6.2 Subjective Measures

In subjective measures, operators perform the task and based on their experience give feedback on the workload measures. The most frequently used self-reports of mental workload are the Subjective Workload Assessment Technique (SWAT) (Reid & Colle, 1988) and the NASA-Task Load Index (Hart & Staveland, 1988). The primary advantages of self-reports are that they are provided directly by the operator involved, they can be collected after the task is done, and they are relatively simple and inexpensive to collect. The disadvantages of self-reports are that the operators are sometimes unaware of their own internal changes, and results can be biased by factors other than workload (e.g. psychosocial environment). These disadvantages can often be overcome if subjective measures are supplemented by one of the other workload measurement approaches.
NASA-TLX was used in this study to measure the subjective workload of participants. This tool defines individual workload scales that are task specific. It consists of six scales; mental demand, physical demand, temporal demand, performance, effort and frustration. Descriptions of each of these scales are given in table 2.2.

<table>
<thead>
<tr>
<th>Workload Component</th>
<th>Endpoints</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental demand (MD)</td>
<td>Low to high</td>
<td>The mental and perceptual activity required by a task</td>
</tr>
<tr>
<td>Physical Demand (PD)</td>
<td>Low to high</td>
<td>The physical activity associated with a task</td>
</tr>
<tr>
<td>Temporal Demand (TD)</td>
<td>Low to high</td>
<td>The time pressure associated with the rate or pace required to complete the task</td>
</tr>
<tr>
<td>Performance</td>
<td>Excellent to poor</td>
<td>The degree of success or satisfaction felt upon the performance or completion of a given task</td>
</tr>
<tr>
<td>Effort</td>
<td>Low to high</td>
<td>The mental and physical work required to perform the task at a certain level</td>
</tr>
<tr>
<td>Frustration</td>
<td>Low to high</td>
<td>Refers to the continuum of stress and/or contentment associated with task completion</td>
</tr>
</tbody>
</table>

2.6.3 Physiological Measures

The human body shows physical reactions to both physically and mentally demanding tasks. Measuring these reactions can be a good indicator of physical and mental work (de Waard, 1996). Many scientists prefer physiological methods of measuring workload over subjective measures because they do not require a direct response from the person and the results do not suffer from subjectivity (Miller, 2001).

Some of the frequently used physiological measures of workload are cardiac activity, respiratory activity, eye activity, speech activity and brain activity. Cardiac activity, so far, is the most common method of measuring workload in driving and aviation experiments (Durantin et al., 2014; Hoover et al., 2012; Roscoe, 1992; Souvestre et al., 2008). Cardiac activity is measured through heart rate, heart rate variability and blood pressure (Hoover et al., 2012).
Respiratory activity involves measuring the number of breaths or the amount of air a person is breathing (Muth et al., 2012). Studies that use eye activity as a measure of workload usually use eye blink rate, horizontal eye movement, and eye closure intervals (Muth et al., 2012). Speech activity measures take pitch, loudness, jitter and shimmer into account (Mendoza & Carballo, 1998). Finally, electroencephalography (EEG) or electrooculography (EOG) are used to measure the electrical activity of the brain (Borghini et al., 2012).

Although the results of physiological measures of workload are reliable (Brookhuis & de Waard, 2010; Reiner & Gelfeld, 2014), they are difficult to implement in the field, due to high costs and obtrusiveness. With the advances in technology and the advent of portable devices like portable heart rate monitoring watches, cardiac measurements are becoming more popular. Cardiac measures are reliable, easy to collect, are unobtrusive and can be collected continuously while the person is performing the job in real world environment (Miller, 2001; Roscoe, 1992; Stuiver et al., 2014).

One of the simplest measures of cardiac activity is heart rate. Generally, as workload increases so does heart rate (Hoover et al., 2012). As an advantage, heart rates are continuously available and the method of collecting heart rate data is not obtrusive. Although it is widely accepted, some researchers criticize this method as it is very sensitive to changes in the psychological state of the person, environmental and emotional changes (Meshkati, 1988; Roscoe, 1992). Another drawback of using heart rate as an indicator of workload is that heart rate does not measure the absolute level of workload. The mean heart rate varies from person to person and when it is used as a measurement, the base measurements should be done first. It is also difficult to differentiate physical workload from mental workload as both cause heart rate to increase (Hoover et al., 2012).
Heart rate variability (HRV) is another cardiac measure of workload. HRV measures inter-beat interval of heartbeats over time (de Waard, 1996). Studies show that during effortful working periods heart rate variability decreases (Stuiver et al., 2014). Heart rate variability is used to evaluate the autonomic nervous system (AS) which acts as a control system for unconscious activities such as heartbeat, digestion, respiratory rate, etc. (Goto et al., 2001; Paritala, 2009; Quintana et al., 2012; Safa-Tisseront et al., 1998). AS is divided into sympathetic and parasympathetic branches. In general, sympathetic activity increases the heart rate and decreases HRV whereas parasympathetic decreases HR and increases HRV (Paritala, 2009). Effortful and mentally demanding tasks are known to be of sympathetic nature.

Scientists use different indices which are extracted from HRV data to measure workload. Generally speaking there are three methods of extracting data. These methods are time domain methods, frequency domain methods and nonlinear methods (Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, 1996).

Time domain methods are the simplest of all and are applied directly to successive intra-beat intervals (NN) (Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, 1996). Table 2.3 shows the common measures used in the time domain method. It is important to note that RMSSD is used as a measure of short time variability (2-5min) whereas SDANN and SDNN require long-term measurements (usually 24 hours). Literature shows that SDNN and RMSSD are two of the most robust time domain measures of workload (Mehler et al., 2011).

In frequency domain methods, a spectrum is calculated from the RR series (where R is a point corresponding to the peak of the QRS complex of the ECG wave; and RR is the interval between successive Rs). Then, this spectrum is divided into three parts. Very low-frequency
(VLF) which ranges from 0 to 0.04 Hz, low-frequency (LF) which ranges from 0.04 to 0.15 Hz, and high-frequency (HF) which ranges from 0.15 to 0.4 Hz (Mehler et al., 2011; Tarvainen et al., 2014; Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, 1996).

Table 2.3 Summary of HRV Time domain parameters (Tarvainen et al., 2014)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN</td>
<td>ms</td>
<td>Standard deviation of all NN intervals</td>
</tr>
<tr>
<td>SDANN</td>
<td>ms</td>
<td>Standard deviation of the averages of NN intervals in all 5-minute segments of the entire recording</td>
</tr>
<tr>
<td>RMSSD</td>
<td>ms</td>
<td>The square root of the mean of the sum of the squares of differences between adjacent NN intervals</td>
</tr>
<tr>
<td>SDNN index</td>
<td>ms</td>
<td>Mean of the standard deviations of all NN intervals for all 5-minute segments of the entire recording</td>
</tr>
<tr>
<td>SDSD</td>
<td>ms</td>
<td>Standard deviation of differences between adjacent NN intervals</td>
</tr>
<tr>
<td>NN50 count</td>
<td></td>
<td>Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording; three variants are possible counting all such NN intervals pairs or only pairs in which the first or the second interval is longer</td>
</tr>
<tr>
<td>pNN50</td>
<td>%</td>
<td>NN50 count divided by the total number of all NN intervals</td>
</tr>
</tbody>
</table>

Table 2.4 shows the common measures used in the frequency domain methods. The power for each variable is the integration of the area under the corresponding band. The low-frequency band is associated with blood pressure control which reflects both sympathetic and parasympathetic activities. The high-frequency band is associated with respiratory sinus arrhythmia which is a parasympathetic activity and the very low-frequency is associated with motor control and temperature control (Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, 1996).
Table 2.4 Summary of HRV frequency domain parameters (Tarvainen et al., 2014)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
<th>Frequency Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-min total power</td>
<td>ms²</td>
<td>The variance of NN intervals over the temporal segment</td>
<td>≈≤0.4 Hz</td>
</tr>
<tr>
<td>VLF</td>
<td>ms²</td>
<td>Power in VLF range</td>
<td>≤0.04 Hz</td>
</tr>
<tr>
<td>LF</td>
<td>ms²</td>
<td>Power in LF range</td>
<td>0.04-0.15 Hz</td>
</tr>
<tr>
<td>LF</td>
<td>Normalized (no unit)</td>
<td>LF power in normalized units LF/(total power−VLF)×100</td>
<td></td>
</tr>
<tr>
<td>HF</td>
<td>ms²</td>
<td>Power in HF range</td>
<td>0.15-0.4 Hz</td>
</tr>
<tr>
<td>HF</td>
<td>Normalized (no unit)</td>
<td>HF power in normalized units HF/(total power−VLF)×100</td>
<td></td>
</tr>
<tr>
<td>LF/HF</td>
<td></td>
<td>Ratio LF [ms²]/HF[ms²]</td>
<td></td>
</tr>
</tbody>
</table>

Studies show that physical workload is linked to HF and mental workload is linked to LF (Paritala, 2009). Low values of HF are the indicator of high physical workload and high values of LF are the indicator of high mental workload. The ratio of LF to HF is defined as an index of parasympathetic and sympathetic balance and can be used as an indicator of mental workload.

As mental workload increases so does LF/HF ratio (Mehler et al., 2011). LF and LF/HF ratio are the most robust measures in the frequency domain measurement of workload (Paritala, 2009). Several studies demonstrated that the normalized value of the LF and HF components could be used to assess sympathetic and parasympathetic activities, respectively (Furlan et al., 2000; Pagani et al., 1997). Time and frequency based domains are considered to be linear methods. Due to the complex behavior of heart and its control system various nonlinear methods have been proposed to fully capture the characteristics of RR intervals. These methods are shown in table 2.5. Detailed information regarding each method is provided by Tarvainen et al. (2014).
Table 2.5 Summary of nonlinear HRV methods (Tarvainen et al., 2014)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD1, SD2</td>
<td>[ms]</td>
<td>Standard deviations of the Poincaré plot</td>
</tr>
<tr>
<td>ApEn</td>
<td>-</td>
<td>Approximate entropy</td>
</tr>
<tr>
<td>SampEn</td>
<td>-</td>
<td>Sample entropy</td>
</tr>
<tr>
<td>D2</td>
<td>-</td>
<td>Correlation dimension</td>
</tr>
<tr>
<td>α1, α2</td>
<td>-</td>
<td>Short-term and long-term fluctuations of detrended fluctuation analysis (DFA)</td>
</tr>
<tr>
<td>Lmean</td>
<td>[beats]</td>
<td>Mean line length of diagonal lines in recurrence plot (RP)</td>
</tr>
<tr>
<td>Lmax</td>
<td>[beats]</td>
<td>Maximum line length of diagonal lines in RP</td>
</tr>
<tr>
<td>REC</td>
<td>[%]</td>
<td>Recurrence rate (percentage of recurrence points in RP)</td>
</tr>
<tr>
<td>DET</td>
<td>[%]</td>
<td>Determinism (percentage of recurrence points which form diagonal lines in RP)</td>
</tr>
<tr>
<td>ShanEn</td>
<td>-</td>
<td>Shannon entropy of diagonal line lengths’ probability distribution</td>
</tr>
</tbody>
</table>

As for physiological variables to study workload, in this study, both time based and frequency based methods to study heart rate variability will be used. The experiments in this study are short term (about 3 minutes). Thus, for time based, RMSSD, and for frequency based, normalized HF, normalized LF and LF/HF ratio will be used.
3 METHODOLOGY

3.1 Study Design

The effect of merge configuration on driver behavior and workload near work zones was measured by using a full-size driving simulator. A 3x2 factorial design with merge type and traffic density as independent variables was used in this study. This research was approved by the LSU Institutional Review Board (IRB) (Appendix 1) and all participants signed a consent form before starting the experiment (Appendix 2). The effect of scenario order was minimized by randomizing the scenario order with a fully counterbalanced Latin Square design.

3.2 Participants

Participants in this study were recruited through convenience sampling from Louisiana State University. 13 female and 17 male students volunteered to participate in this study. The criteria for inclusion were having a valid driving license, not being pregnant, and not being prone to motion sickness. The age of participants ranged from 20 to 35 years with a median of 22.5. On average, participants had 7 years of driving experience with an average of 12,350 miles driven per year. The self-reported questions regarding driving experience showed that out of 30 participants, two were involved in an accident previously and four had violated driving laws resulting in a ticket in the past 12 months.

A priori power analyses was conducted to estimate the number of required participants for this study by using G*power 3.1.9 (Faul et al., 2007). In a priori power analyses, sample size N is computed as a function of the required power level (1-β), the pre-specified significance level, and the population effect size to be detected with probability 1-β (Cohen, 1988). The priori power analysis showed 30 participants as the recommended sample size for a MANOVA.
repeated measures, between groups with a medium effect size ($f^2 = 0.25$), $\alpha=0.05$, power level $(1-\beta ) = 0.90$.

![G*Power 3.1.9](image)

Figure 3.1 Result of G*power estimates for the number of required participants in this study

3.3 Tools

A full-size driving simulator (Realtime Technologies Inc., Royal Oak, MI) as shown in figure 3.2 was used in this study to simulate driving through a construction zone. The simulator is a full size passenger car on a one degree-of-freedom motion base, providing realistic motion cues to the driver, and is surrounded by four screens showing front, rear, left and right views.
with 1680 x 1050 screen resolution. The core simulator and visuals subsystems operate at a 60 Hz update rate, supporting smooth graphics presentation and rapid system response in complex driving environments (Realtime Technologies Inc, 2014). The driving simulator has demonstrated measured latency of less than 50 ms from step input on the host to visuals output. The side-mirrors consist of two LCDs which show the rear view of the road. There are three cameras inside and one camera outside the car to record eye movement, foot position on accelerator and gas pedals, steering wheel and ambient traffic flow.

Figure 3.2 View of driving simulator

A Polar RS800 heart rate monitoring watch (Polar Electro, Kempele, Finland) was used to collect heart rate data. The watch comes with a sensor belt which is worn around the chest. This belt transmits heart rate data to the watch through infrared connection. After the experiment, the data was transferred to a computer through a USB adaptor for further analysis. Figure 3.3 shows the Polar RS800 and its accessories.
3.4 Experimental Model

Six driving scenarios were designed based on an interstate highway driving environment (two work zone configurations plus a control group × two levels of traffic density). The two types of work zone configurations selected for this study were conventional merge (CLM) and joint merge (JLM). An interstate highway with no work zone was used as a control group.
The driving scenarios were comprised of a 3.7 miles two lane highway with a construction zone located on the right lane. There were no traffic lights, yield signs or stop signs. The speed limits prior and after the work zone was 70 mph. The length of the road prior to the advance warning zone was 2 miles. This length was enough for cars to reach the posted speed of 70 mph. Work zones were designed with posted speed limits of 50 mph. The signs presented posted speed limits and distances in U.S. customary units (MPH and miles). A large stop sign was placed at the end of the simulated highway and participants were asked to stop before this sign at the end of the simulation. The simulator provides traffic densities ranging from 0 to 50 vehicles–per–mile. The average density was 10 vpm and 50 vpm for low and high traffic density, respectively. The ambient traffic in traffic simulator is responsive to participants’ driving behavior. The speed of other vehicles in the simulator was set to 70 mph before the speed limit sign and 50 mph after the speed limit sign. The merging behavior by other cars in the simulator is similar to that of in the real world. The cars in the closed lanes merged into the open lane in an alternating order and usually there was not a big gap between the cars.

The CLM, JLM and late merge layouts were divided into five different zones as shown in Figures 3.5-6. These zones are (1) advance warning zone, (2) transition zone, (3) buffer, (4) work zone and (5) termination zone. The advance warning zone is typically a mile long and is primarily used to inform the motorists of what to expect ahead as they approach the work zone area. When redirection of the driver's normal path is required, traffic must be channelized from the normal path to a new path. This redirection is done in the transition area. The buffer space is an optional feature in the activity area that separates traffic flow from the work activity or a potentially hazardous area and provides recovery space for an errant vehicle. The work zone is an area of roadway where the work takes place. It is composed of the work space and the traffic
space, and may contain one or more buffer spaces. Traffic returns to the normal traffic path in the termination area. The termination area extends from the downstream end of the work area to the END ROAD WORK signs, if posted.

Figure 3.5 Conventional merge layout with right lane closure

Figure 3.6 Joint lane merge configuration
3.5 Independent and Dependent Variables

The independent variables used in this study are:

- Merge configuration (CLM, JLM and control group with no work zone)
- Traffic density (high or low)

The dependent variables used in this study are categorized into three groups, physiological measures, subjective workload ratings, and performance variables, as shown in Tables 3.1, 3.2 and 3.3.

| Table 3.1 Physiological measures of workload |
|-----------------|---------|-------------------------------------------------|
| Variable | Unit | Description |
| RMSSD | ms | The square root of the mean of the sum of the squares of differences between adjacent NN intervals |
| LF | n.u. | Power in LF range in normalized units |
| HF | n.u. | Power in HF range in normalized units |
| LF/HF | - | Ratio between LF and HF band powers |

The LF and HF in normalized units were calculated by equations 3.1 and 3.2 (Tarvainen et al., 2014);

\[
LF(n.u.) = \frac{LF (ms^2)}{[Total \ Power (ms^2) - VLF (ms^2)]} \tag{3.1}
\]

\[
HF(n.u.) = \frac{HF (ms^2)}{[Total \ Power (ms^2) - VLF (ms^2)]} \tag{3.2}
\]

where LF (ms\(^2\)) and HF (ms\(^2\)) are absolute powers of low frequency and high frequency bands.

Total power is the total spectral power and VLF (ms\(^2\)) is the absolute power for very low frequency band. All these variables are calculated automatically by Kubios (Biosignal Analysis and Medical Imaging Group, Finland), when NN interval data from the heart monitoring watch is fed into the software.
Subjective workload was measured by using NASA-TLX (Hart & Staveland, 1988) (Appendix 4). NASA TLX is a multidimensional test which measures mental demand, physical demand, temporal workload, frustration, performance, and effort.

Table 3.2 Subjective workload variables from the NASA-TLX

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental demand</td>
<td>%</td>
<td>The mental and perceptual activity required by a task</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>%</td>
<td>The physical activity associated with a task</td>
</tr>
<tr>
<td>Temporal Demand</td>
<td>%</td>
<td>The time pressure associated with the rate or pace required to complete the task</td>
</tr>
<tr>
<td>Performance</td>
<td>%</td>
<td>The degree of success or satisfaction felt upon the performance or completion of a given task</td>
</tr>
<tr>
<td>Effort</td>
<td>%</td>
<td>The mental and physical work required to perform the task at a certain level</td>
</tr>
<tr>
<td>Frustration</td>
<td>%</td>
<td>The continuum of stress or contentment associated with task completion</td>
</tr>
<tr>
<td>Total Workload</td>
<td>%</td>
<td>Weighted sum of all TLX components</td>
</tr>
</tbody>
</table>

To measure subjective workload, each participant was given a workload rating sheet (Appendix 3) after doing a task. The participants should decide how much workload he/she experienced while performing the task. The magnitude of workload is marked by putting a vertical line on a 12cm visual-analog scale. The distance from the left end of the 12 cm line to the marking represents the rating for that scale (Hart & Staveland, 1988). For each task a separate workload rating sheet was completed by the participants. At the end of the experiment a scale comparison sheet was given to the participant to do a pairwise comparison between the six scales. The goal was to find out which of the scales had a relative dominance over the other scales. This process was used to calculate the weight of each scale. At the end each rating was multiplied by its corresponding scale’s weight. Total workload can be also calculated as the sum of all scales.
Driver’s hand position on the steering wheel can reflect the perceived risk of the road context (Walton & Thomas, 2005). Furthermore, hand positions can reflect a response to the demand of the situation and could be more closely related to mental workload and the need to control steering rather than a response to perceived risk (de Waard et al., 2010; Lewis-Evans & Rothengatter, 2009). In this study, steer variability is the variability in the rotation of steering wheel to the left or right. The other performance variables selected for this study were variability in the force exerted on the brake pedal and variability in speed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed Variability</td>
<td>m/s</td>
<td>Standard deviation of vehicle’s speed</td>
</tr>
<tr>
<td>Steer Variability</td>
<td>degrees</td>
<td>Standard deviation of angle of the steering wheel</td>
</tr>
<tr>
<td>Brake force variability</td>
<td>N</td>
<td>Standard deviation of the force exerted on the brake pedal</td>
</tr>
</tbody>
</table>

### 3.6 Procedure

Before starting the driving scenarios, participants were briefed on the purpose and risks of the study and instructed how to complete the scenarios. After briefing, participants wore the RS800 transmitter belt around their chest. Then, a set of forms were given to each participant including the informed consent form and demographic information form. The experimenter explained how to use the NASA-TLX response sheet and the six subcomponents of workload. On average, filling all the forms and questionnaires took about 5 minutes. As each participant was filling the forms, their heart rate was recorded which later was used as a baseline heart rate. The base line heart rate was collected for 3 minutes prior to the start of the experiment. To ensure participants felt well prior to the start of experiment and to provide a baseline to compare their health symptoms after completing scenarios, participants completed a motion sickness
questionnaire (Gianaros et al., 2001) (Appendix 5). Participants then were allowed to familiarize themselves with the simulator by driving on a test road consisting of a two mile interstate highway with a work zone on the left lane. After completing the test, participants were asked to rank the importance of each subcomponent of the NASA-TLX through the set of pair-wise comparisons.

During the experiment, the researcher sat outside the simulator at a desktop station. Participants were asked to obey traffic rules and posted speed limits. They were not given any information on how and when to merge because that would create bias in the experiment. Furthermore, they did not know whether the work zone was on the left or right side of the road. However, the start position of each experiment was on the right shoulder of the road. The length of each drive was approximately two to three minutes. Heart rate data was collected for each scenario separately. At the beginning of each scenario, the experimenter pressed the record button on the Polar watch to start recording heart rate data and at the end pressed the stop button to finish recording. Throughout the experiment the radio was off but the noise from ambient traffic was playing through several speakers around the simulator. At the completion of each scenario, there was a 2-minute break, during which participant completed the NASA TLX rating sheet. Participants completed six scenarios, and the motion sickness questionnaire (Appendix 5) was repeated after every three trials to ensure participants were not having motion sickness.

3.7 Data Analysis

3.7.1 Pre-processing

After the six scenarios were complete, heart rate data was transferred to a computer through the USB adapter. PolarPro Trainer 5, which is software that comes with Polar RS800, was used for downloading the data from the watch. PolarPro 5 is capable of exporting inter-beat
beats interval (RR) to a text file. Figure 3.8 shows a sample of exported RR data. This text file is then imported to Kubios HRV for calculating HRV indices.

![Test - Notepad](image)

Figure 3.7 Sample exported RR data from PolarPro 5

Figure 3.9 shows the Kubios HRV interface. Pallet 1 gives general information about the collected data such as duration, start and end point etc. Pallet 2 is used to define frequency bands as discussed in chapter 3.

The third window plots RR values over time. All the HRV indices are provided in part 4. Kubios HRV calculate time domain, frequency domain and nonlinear methods. These calculations are performed automatically when the RR text file is loaded into the program. Figure 3.9-10 show the snap shots of the measures provided in each method. Selected HRV measures for each participant were calculated and used for statistical analysis.
Figure 3.8 Example Kubios HRV interface

Figure 3.9 Example Kubios HRV time domain measures
3.7.2 Statistical pre-processing

All statistical analyses were done by using SPSS statistical package version 21 (IBM Corp, 2012). All tests were done at alpha level of 0.05. In the first step, the collected data were screened for any outliers. Appendix 6 shows the histograms for physiological measures of workload. The outlier labeling rule proposed by Hoaglin et al. (1986) and Hoaglin and Iglewicz (1987), which is based on multiplying the Interquartile Range (IQR) by a factor of 2.2 was used in this study to detect the outliers. Based on this rule any data point that falls beyond the \([Q_1 - 2.2IQR], (Q_3 + 2.2IQR)\] limits is considered an outlier. Accordingly, five data points in RMSSD and three data points in LF/HF ratio were identified as outliers and removed from the dataset. In order to increase the robustness and power of multivariate test, sample sizes should be equal. Thus, the deleted outliers were replaced by using Expectation-Maximization (E-M) algorithm (Little & Rubin, 2002) embedded in SPSS. The E-M imputation algorithm starts by estimating the expected values of missing data from observed data and then repeats the process using both the observed data and the estimated missing values. The process repeats until the values stabilize (Allison, 2002).
Tests of normality (Kolmogorov-Smirnov) and homogeneity of variance (Levene’s test of equality of variances) were performed prior to conducting any inferential statistics. Table 4.7 shows the variables which failed the normality and homogeneity of variance tests. Except for normalized LF and HF, frustration and total workload, the rest of variables were not normal. Furthermore, mental demand along with performance variables did not pass the homogeneity of variance test. For all the measures that did not achieve a satisfactory level of normality, Johnson transformation with Bounced System (SB) method was applied using Minitab 16 (Minitab Inc, 2010). This transformation which is proposed by Yeo and Johnson (2000) is very effective in correcting skewness as well as heavy tails in variables that are both positive and negative. Equation 1 was used to transform non-normal data into normal data.

\[ y = \gamma + \eta \ln \left( \frac{x - \varepsilon}{\lambda + \varepsilon - \chi} \right) \]  

(3.1)

Where \( y \) is the transformed value.

\( \gamma \) is the shape 1 parameter.

\( \eta \) is the shape 2 parameter.

\( \varepsilon \) is the location parameter.

\( \lambda \) is the scale parameter.
Table 3.4 list of transformed variables

<table>
<thead>
<tr>
<th>Variable Category</th>
<th>Variable Name</th>
<th>Before Transformation</th>
<th>After Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Normality Test (p value)</td>
<td>Homogeneity of Variance (p value)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological measures of workload</td>
<td>RMSSD</td>
<td>X (0.00)</td>
<td>✓ (0.067)</td>
</tr>
<tr>
<td></td>
<td>Normalized LF</td>
<td>✓ (0.2)</td>
<td>✓ (0.231)</td>
</tr>
<tr>
<td></td>
<td>Normalized HF</td>
<td>✓ (0.2)</td>
<td>✓ (0.106)</td>
</tr>
<tr>
<td></td>
<td>LF/HF</td>
<td>X (&lt;0.001)</td>
<td>✓ (0.906)</td>
</tr>
<tr>
<td>Subjective workload</td>
<td>Mental</td>
<td>X (&lt;0.001)</td>
<td>X (0.025)</td>
</tr>
<tr>
<td></td>
<td>Physical</td>
<td>X (&lt;0.001)</td>
<td>✓ (0.22)</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
<td>X (&lt;0.001)</td>
<td>✓ (0.049)</td>
</tr>
<tr>
<td></td>
<td>Performance</td>
<td>X (0.003)</td>
<td>✓ (0.984)</td>
</tr>
<tr>
<td></td>
<td>Effort</td>
<td>X (0.005)</td>
<td>✓ (0.514)</td>
</tr>
<tr>
<td></td>
<td>Frustration</td>
<td>✓ (0.2)</td>
<td>✓ (0.851)</td>
</tr>
<tr>
<td></td>
<td>Total Workload</td>
<td>✓ (0.2)</td>
<td>✓ (0.457)</td>
</tr>
<tr>
<td>Performance Variables</td>
<td>Steer variability</td>
<td>X (&lt;0.001)</td>
<td>X (&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Brake variability</td>
<td>X (&lt;0.001)</td>
<td>X (&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Speed Variability</td>
<td>X (&lt;0.001)</td>
<td>X (&lt;0.001)</td>
</tr>
</tbody>
</table>

After transformation the number of non-normal variables reduced from ten to two. Two variables violated the assumptions of homogeneity. However, since MANOVA is a robust test, they were in the model.

3.7.3 Hypothesis Testing

Statistical analysis tests the hypothesis that physiological measures of workload, subjective workload and performance variables are influenced by work zone configuration, traffic density or the interaction between these two. To test these hypotheses a multivariate analysis of variance was conducted with the probability of type I error of 0.05. This study tested the following hypothesis:
• Hypothesis 1:
  o $H_0$: There is no difference in drivers’ physiological workload at the presence of work zones in different traffic densities.
  o $H_a$: There is at least a difference in drivers’ physiological workload at the presence of work zones in different traffic densities.

• Hypothesis 2:
  o $H_0$: There is no difference in drivers’ subjective workload at the presence of work zones in different traffic densities.
  o $H_a$: There is a difference in at least one measure of drivers’ subjective workload at the presence of work zones in different traffic densities

• Hypothesis 3:
  o $H_0$: There is no association between subjective workload, physiological workload and performance variables
  o $H_a$: There is at least an association between subjective workload, physiological workload and performance variables

To check the effect of order of scenarios on dependent variables, a Multivariate test with all 14 dependent variables was conducted and the order was entered into the model as the only fixed effect. The result was not significant, indicating that the carryover effect was minimal, Wilks’ $\lambda=0.039$, $F(70, 56.45)=0.788$, $p=0.829$, $\eta^2_{partial}=0.447$, power=0.706.

A two-way MANOVA revealed significant multivariate main effects for driving scenarios, Wilks’ $\lambda=0.666$, $F(28, 322)=2.58$, $p<0.001$, $\eta^2_{partial}=0.18$, power>0.99 and traffic density Wilks’ $\lambda=0.443$, $F(14, 161)=14.43$, $p<0.001$, $\eta^2_{partial}=0.56$, power>0.99. The driving scenarios and traffic interaction were also found to be significant, Wilks’ $\lambda=0.583$, $F(28,
322) = 3.56, $p < 0.001$, $\eta^2_{\text{partial}} = 0.236$, power $> 0.99$. Given the significance of the overall test, univariate main effects were examined.
4 RESULTS

4.1 Physiological Measures of Workload

Table 4.1 and figure 4.1 show the descriptive statistics for physiological measures of workload in different driving scenarios. RMSSD, which is the measure of variation in heartbeat intervals and is used to measure physical workload was the highest in JLM and the lowest in CLM, though the difference (6%) was not statistically significant, $F(2, 174)=0.18,$ $p=0.839, \eta^2_{partial}=0.002, \text{power}=0.077$.

Table 4.1 Summary statistics (means and standard deviations) for physiological measures of workload at rest and in three driving scenarios (N=60*)

<table>
<thead>
<tr>
<th>HRV Variables</th>
<th>Rest</th>
<th>Driving Scenarios</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>RMSSD</td>
<td>69.04</td>
<td>3.66</td>
<td>49.42</td>
<td>26.90</td>
<td>48.59</td>
<td>24.96</td>
</tr>
<tr>
<td>Normalized LF</td>
<td>0.58</td>
<td>0.08</td>
<td>0.63</td>
<td>0.14</td>
<td>0.62</td>
<td>0.14</td>
</tr>
<tr>
<td>Normalized HF</td>
<td>0.36</td>
<td>0.10</td>
<td>0.38</td>
<td>0.13</td>
<td>0.37</td>
<td>0.13</td>
</tr>
<tr>
<td>LF/HF</td>
<td>1.76</td>
<td>0.68</td>
<td>2.12</td>
<td>1.74</td>
<td>2.09</td>
<td>1.48</td>
</tr>
</tbody>
</table>

* for the rest period N =30

LF, which is used to measure mental demand, was almost the same in the roads with a work zone compared to the roads without a work zone, $F(2, 174 )=0.58, p=0.557$, $\eta^2_{partial}=0.007, \text{power}=0.147$. HF, which is used as an indicator of physical demand, was not significantly different in driving scenarios, $F(2, 174 )=0.12, p=0.887, \eta^2_{partial}=0.001$, power=0.068. Lower values of HF are an indicator of more physical demand. The ratio of LF to HF, which is a measure of mental demand was also not statistically different in different driving scenarios $F(2, 174 )=0.18, p=0.832, \eta^2_{partial}=0.002, \text{power}=0.078$. The highest value of LF/HF ratio belonged to the control group.
Figure 4.1 Mean physiological measures of workload in different driving scenarios (a) RMSSD; (b) LF; (c) HF; (d) LF/HF

Table 4.2 and figure 4.2 show the descriptive statistics for physiological measures of workload in different traffic densities. RMSSD in high traffic density was 9.6% lower than that of in low traffic density. This indicates that driving in high traffic density was 9.6% more physically demanding than driving in low traffic density. However, this difference was not statistically significant, $F(1, 174) = 0.495, p = 0.482, \eta^2_{\text{partial}} = 0.003$, power = 0.108. Although the time domain measure of heart rate variability changed in different traffic conditions, LF, $F(1, 174) = 0.026, p = 0.873, \eta^2_{\text{partial}} = 0.001$, power = 0.053, and HF, $F(1, 174) = 0.014, p = 0.906,
\( \eta_{partial}^2 = 0.001, \) power=0.052, were unaffected by the presence of traffic. LF/HF ratio suggests that driving in high traffic density was only 3% less mentally demanding than driving in low traffic density, \( F(1, 174) = 0.001, p=0.970, \eta_{partial}^2 = 0.001, \) power=0.05. No significant interaction between driving scenarios and traffic density was observed for physiological measures of workload (p>0.05)

Table 4.2 Summary statistics (means and standard deviations) for physiological measures of workload at rest and in different traffic densities (N = 90)

<table>
<thead>
<tr>
<th>HRV Variables</th>
<th>Rest Mean</th>
<th>Std.</th>
<th>High Traffic Density Mean</th>
<th>Std.</th>
<th>Low Traffic Density Mean</th>
<th>Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSSD</td>
<td>69.04</td>
<td>3.66</td>
<td>47.45</td>
<td>23.11</td>
<td>52.54</td>
<td>30.61</td>
</tr>
<tr>
<td>Normalized LF</td>
<td>0.58</td>
<td>0.08</td>
<td>0.62</td>
<td>0.14</td>
<td>0.62</td>
<td>0.15</td>
</tr>
<tr>
<td>Normalized HF</td>
<td>0.36</td>
<td>0.10</td>
<td>0.38</td>
<td>0.12</td>
<td>0.38</td>
<td>0.15</td>
</tr>
<tr>
<td>LF/HF</td>
<td>1.76</td>
<td>0.68</td>
<td>2.04</td>
<td>1.57</td>
<td>2.1</td>
<td>1.52</td>
</tr>
</tbody>
</table>

* for the rest period N =30

Figure 4.2 Mean physiological measures of workload in different traffic densities (a) RMSSD; (b) LF
4.2 Subjective Measures of Workload

Table 4.3 and figure 4.3 show the descriptive statistics associated with the TLX components in different driving scenarios. Results showed that driving scenarios was a significant main effect for mental demand, $F(2,174)=0.593, p=0.003, \eta^2_{partial}=0.064$, power=0.874; temporal demand, $F(2,74)=3.415, p=0.035, \eta^2=0.038$, power=0.636; performance, $F(2,174)=3.415, p=0.035, \eta^2=0.038$, power=0.636; and effort, $F(2,174)=4.361, p=0.014, \eta^2=0.048$, power=0.749. A post-hoc Tukey test showed CLM and JLM did not differ significantly from each other ($p > 0.05$). However, CLM was significantly higher than the control group with respect to mental demand ($p=0.003$), temporal demand ($p=0.029$), effort ($p=0.048$) and total workload ($p=0.011$). Only effort in JLM was significantly higher than that of in the control group ($p=0.020$).
Table 4.3 Summary statistics (means and standard deviations) for NASA TLX components in three driving scenarios (N=60)

<table>
<thead>
<tr>
<th>Workload Subcomponent</th>
<th>Control Group</th>
<th>CLM</th>
<th>JLM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
</tr>
<tr>
<td>Mental</td>
<td>48.00</td>
<td>19.94</td>
<td>59.35</td>
</tr>
<tr>
<td>Physical</td>
<td>26.20</td>
<td>19.83</td>
<td>30.55</td>
</tr>
<tr>
<td>Temporal</td>
<td>41.10</td>
<td>25.16</td>
<td>50.78</td>
</tr>
<tr>
<td>Performance</td>
<td>31.53</td>
<td>22.21</td>
<td>32.07</td>
</tr>
<tr>
<td>Effort</td>
<td>39.87</td>
<td>21.33</td>
<td>49.07</td>
</tr>
<tr>
<td>Frustration</td>
<td>46.62</td>
<td>25.23</td>
<td>42.05</td>
</tr>
<tr>
<td>Total Workload</td>
<td>41.25</td>
<td>17.02</td>
<td>48.67</td>
</tr>
</tbody>
</table>

Figure 4.3 Mean percent subjective workload for different driving scenarios
Table 4.4 and figure 4.4 show the comparison of NASA TLX subcomponents in two traffic densities. The results showed that except for performance which was not significantly different in two traffic densities, $F(1,174)=1.963$, $p=0.163$, $\eta^2_{\text{partial}}=0.011$, power=0.286, the remaining TLX components were significantly higher in the high traffic density ($p<0.05$).

Results indicated that drivers in high traffic density experienced 22% more mental demand, $F(1,174)=21.271$, $p<0.001$, $\eta^2_{\text{partial}}=0.109$ power=0.996; 25% more physical demand, $F(1,174)=4.325$, $p=0.039$, $\eta^2_{\text{partial}}=0.024$, power=0.543; 24% more temporal demand, $F(1,174)=12.068$, $p=0.001$, $\eta^2_{\text{partial}}=0.065$, power=0.993; 32% more effort, $F(1,174)=17.80$, $p<0.001$, $\eta^2_{\text{partial}}=0.093$, power=0.987; and 52% more frustration, $F(1,174)=30.345$, $p<0.001$, $\eta^2_{\text{partial}}=0.148$, power=0.999. Overall, the total workload for driving in high traffic density was 27% more than that of in low traffic density, $F(2,174)=27.537$, $p<0.001$, $\eta^2_{\text{partial}}=0.137$, power=0.999. There was no significant driving scenario with traffic density interaction for subjective measures of workload.

Table 4.4 Comparison of NASA TLX components (means and standard deviations) in two traffic density levels (N=90)

<table>
<thead>
<tr>
<th>Variable</th>
<th>High Traffic Density</th>
<th>Low Traffic Density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>Mental</td>
<td>59.42</td>
<td>18.23</td>
</tr>
<tr>
<td>Physical</td>
<td>32.50</td>
<td>20.10</td>
</tr>
<tr>
<td>Temporal</td>
<td>50.63</td>
<td>20.71</td>
</tr>
<tr>
<td>Performance</td>
<td>34.70</td>
<td>20.72</td>
</tr>
<tr>
<td>Effort</td>
<td>52.79</td>
<td>20.87</td>
</tr>
<tr>
<td>Frustration</td>
<td>53.69</td>
<td>23.25</td>
</tr>
<tr>
<td>Total Workload</td>
<td>50.93</td>
<td>13.42</td>
</tr>
</tbody>
</table>
4.3 Performance Variables

Table 4.5 and figure 4.5 show the descriptive statistics for the performance variables used in this study. Results showed that driving scenarios was a significant main effect for steer variability, $F(2, 174) = 6.262, p = 0.002, \eta^2_{partial} = 0.067$, power = 0.891, and speed, $F(2, 174) = 7.440, p = 0.001, \eta^2_{partial} = 0.079$, power = 0.938. No significant differences were observed in brake variability between different scenarios, $F(2, 174) = 1.387, p = 0.253, \eta^2_{partial} = 0.016$, power = 0.295. A Post-hoc Tukey test for steer variability showed that there was no difference in steer variability between CLM with JLM ($p = 0.381$) and CLM with the control group ($p = 0.078$) but JLM was significantly higher than the control group ($p = 0.002$). In terms of speed variability, CLM ($p < 0.002$) and JLM ($p < 0.005$) were both significantly higher than the control group but they were not significantly different from each other ($p = 0.945$).
Table 4.5 Descriptive statistics (means and standard deviations) for performance variables in three driving scenarios (N=60)

<table>
<thead>
<tr>
<th>Performance Variable</th>
<th>Control Mean</th>
<th>Control Std.</th>
<th>CLM Mean</th>
<th>CLM Std.</th>
<th>JLM Mean</th>
<th>JLM Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steer Variability (degrees)</td>
<td>43.69</td>
<td>79.52</td>
<td>37.24</td>
<td>28.49</td>
<td>45.68</td>
<td>49.49</td>
</tr>
<tr>
<td>Brake Variability (N)</td>
<td>2.83</td>
<td>4.70</td>
<td>2.58</td>
<td>3.34</td>
<td>1.90</td>
<td>3.58</td>
</tr>
<tr>
<td>Speed Variability (m/s)</td>
<td>1.36</td>
<td>0.91</td>
<td>2.14</td>
<td>1.44</td>
<td>2.02</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Figure 4.5 Mean variability for performance variables for different driving scenarios

Table 4.6 and figure 4.6 show that steer variability and brake variability were higher in high traffic density. On average, the variability in turning the steering wheel in high traffic density was 21.27 degrees more which is 67.3% more than that of in low traffic density, $F(1,174) = 7.728, p=0.006, \eta^2_{partial}=0.043$, power=0.789. Brake variability in high traffic density was 3.8 times more, $F(1,174) = 40.636, p<0.001, \eta^2_{partial}=0.189$, power>0.99. In terms of speed
variability, driving in low traffic density resulted in 15% more variability in speed,

\[ F(1,174) = 42.034, \ p < 0.001, \ \eta^2_{\text{partial}} = 0.195, \ \text{power} > 0.99. \]

Table 4.6 Comparison of performance variables (means and standard deviations) in two traffic density levels (N=90)

<table>
<thead>
<tr>
<th></th>
<th>High Traffic Density</th>
<th>Low Traffic Density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>Steer variability (degrees)</td>
<td>52.84</td>
<td>73.81</td>
</tr>
<tr>
<td>Brake variability (N)</td>
<td>3.87</td>
<td>4.57</td>
</tr>
<tr>
<td>Speed Variability (m/s)</td>
<td>1.27</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Figure 4.6 Mean variability for performance variables in different traffic densities

Univariate tests revealed significant interaction between driving scenarios and traffic densities for brake variability, \( F(2,174) = 9.983, \ p < 0.001, \ \eta^2_{\text{partial}} = 0.103, \ \text{power} = 0.984, \) and speed variability, \( F(2,174) = 29.704, \ p < 0.001, \ \eta^2_{\text{partial}} = 0.225, \ \text{power} > 0.99. \) With respect to brake variability figure 4.1 shows that brake variability remains constant for CLM regardless of traffic density, but increases substantially in the JLM when traffic density increases. Figure 4.2
shows that CLM and JLM behave almost identically, but are very different from control group. Control group had the highest speed variability in high traffic density and CLM had the highest in low traffic density. The lowest speed variability in high and low traffic density were observed in the JLM and control group, respectively.

Figure 4.7 Driving scenario-traffic density interaction graph for brake variability

Figure 4.8 Driving scenario-traffic density interaction graph for speed variability
4.4 Relationship between Physiological Measures of Workload, Subjective Workload and Performance Variables

Pearson correlation coefficients were calculated in order to determine if there were any relationships between physiological measures of workload, subjective measures of workload and performance variables (Table 4.7). Results indicated RMSSD correlated weakly but significantly with all NASA TLX components except for performance. There was no linear relationship between performance and RMSSD.

LF correlated weakly with subjective mental demand \((r= -0.214)\) and total workload \((r= -0.156)\). HF correlated positively with mental demand \((r= 0.213)\) and total workload \((r= 0.152)\). LF/HF ratio correlated weakly with total workload \((r= -0.164)\). No significant correlation was found between physiological measures of workload and performance variables.

Steer variability correlated weakly with temporal demand \((r =0.196)\), performance \((r=0.158)\), and total workload \((r =0.193)\). Brake variability correlated positively with mental demand \((r =0.255)\), temporal demand \((r =0.273)\), effort \((r =0.263)\), frustration \((r =0.320)\) and total workload \((r =0.352)\). No linear relationship was found between subjective workload and speed variability.
Table 4.7 Pearson correlation matrix for physiological measures of workload, subjective workload and performance variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 RMSSD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 LF (Normalized)</td>
<td>-.437**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 HF (Normalized)</td>
<td>.487**</td>
<td>-.858**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 LF/HF</td>
<td>-.492**</td>
<td>.944**</td>
<td>-.972**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Mental</td>
<td>.318**</td>
<td>-.214**</td>
<td>.213**</td>
<td>-.229**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Physical</td>
<td>.246**</td>
<td>-.120</td>
<td>.132</td>
<td>-.133</td>
<td>.320**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Temporal</td>
<td>.215**</td>
<td>-.061</td>
<td>.064</td>
<td>-.079</td>
<td>.445**</td>
<td>.268**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Performance</td>
<td>.042</td>
<td>-.079</td>
<td>.095</td>
<td>-.084</td>
<td>.115</td>
<td>.164*</td>
<td>.175*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Effort</td>
<td>.276**</td>
<td>-.145</td>
<td>.122</td>
<td>-.142</td>
<td>.558**</td>
<td>.237**</td>
<td>.445**</td>
<td>.359**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Frustration</td>
<td>.211**</td>
<td>.001</td>
<td>-.001</td>
<td>-.007</td>
<td>.326**</td>
<td>.133</td>
<td>.430**</td>
<td>.183*</td>
<td>.342**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Total Workload</td>
<td>.312**</td>
<td>-.156*</td>
<td>.152*</td>
<td>-.164*</td>
<td>.741**</td>
<td>.419**</td>
<td>.692**</td>
<td>.540**</td>
<td>.789**</td>
<td>.521**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Steer Variability</td>
<td>-.027</td>
<td>.168*</td>
<td>-.158*</td>
<td>.161*</td>
<td>.127</td>
<td>.111</td>
<td>.196**</td>
<td>.158*</td>
<td>.113</td>
<td>.144</td>
<td>.193**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Brake Variability</td>
<td>.088</td>
<td>.054</td>
<td>-.077</td>
<td>.049</td>
<td>.255**</td>
<td>.081</td>
<td>.273**</td>
<td>.109</td>
<td>.263**</td>
<td>.320**</td>
<td>.352**</td>
<td>.370**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14 Speed Variability</td>
<td>.004</td>
<td>.039</td>
<td>-.067</td>
<td>.060</td>
<td>-.069</td>
<td>-.049</td>
<td>-.012</td>
<td>-.054</td>
<td>-.048</td>
<td>-.103</td>
<td>-.075</td>
<td>.231**</td>
<td>.261**</td>
<td>1</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).
5 DISCUSSION

The present study was carried out to first, evaluate the impacts of different work zone configurations and traffic density on physiological measures of workload, subjective workload and performance variables; and second, to examine the relationship between physiological measures of workload, subjective workload and performance variables.

Variation in beat-to-beat intervals of heart (NN) is a physiological phenomenon and can be caused by physical activity and stress. Madden and Savard (1995) and Hjortskov et al. (2004) associated low heart rate variability to mental and physical stress. Research shows that an increase in LF is associated with sympathetic activity (mental stress) and reduction in HF is linked with parasympathetic activity (physical stress) (Kamath & Fallen, 1993; Wang et al., 2005). Moreover, an increase in the LF/HF ratio is an indicator of increased mental workload (Durantin et al., 2014; Hjortskov et al., 2004).

Comparing the measures of HRV at rest with those collected during the driving simulation suggests that participants, as compared to the rest period, experience more workload when driving in the simulator. Results showed lower RMSSD during the driving scenarios suggesting lower heart rate variability in driving.

Although there were some differences between control group, CLM and JLM in terms of measures of HRV, none of the measures of RMSSD, LF, HF and LF/HF were significantly different from each other. Similar inconsistencies were found in the literature on HRV measures of mental and physical workload. In a simulation study, Veltman and Gaillard (1996) found that in effortful scenarios, HRV measures were unaffected. They found that this insensitivity was caused by the respiratory activity as a confounding factor. When the respiratory frequency
decreased, HRV increased and vice versa. Stuiver et al. (2014) studied the effects of short increases in task demand on heart rate and blood pressure as indicators of mental effort in a driving simulator study. Their study was comprised of six sessions of driving in a driving simulator in two levels of traffic density. The increase in the task demand was simulated by driving through fog for 40 seconds. The results showed lower blood pressure variability in higher traffic density. Heart rate variability and blood pressure variability measures decreased during driving in fog in the low traffic condition, indicating increased effort investment during fog in this condition. Mulder et al. (1992), on the other hand, showed that in long tasks, the initial HRV effects disappeared after 10 to 20 min of task performance, while BP remained high after the initial effects. The authors concluded that these effects were directly related to short term blood pressure control. Stuiver et al. (2014) suggested that in case of increased task demand, cardiovascular responses to increased mental workload can be either a continuing rise in heart rate (initial reaction) or a decrease in heart rate (regulation effect). The combination of these two effects may largely explain the mixed results on heart rate and heart rate variability measures that are found in some studies (Mulder et al., 1992; Porges & Byrne, 1992; Sirevaag et al., 1993; Veltman & Gaillard, 1996; Wilson, 1992).

LF and HF showed no sensitivity to the changes in the driving scenarios, and they were almost equal in all three driving scenarios. Despite no statistical differences, RMSSD and LF/HF ratio were more sensitive to the changes in the work zones than LF and HF. We can attribute the insensitivity of LF and HF measures to changes in the work zone configurations to several factors. Since all six scenarios were done in one session (approximately 45 minutes), and between each scenario participants had a short break to fill TLX questionnaire, participants may have acclimatized to the task demands, similar to findings in the Mulder et al. (1992) experiment.
described previously. Moreover, low values of partial eta squared ($\eta^2_{\text{partial}}$), which indicate the low percentage of variance in each of the effects (or interaction) and its associated error that is accounted for by that effect (or interaction), suggest that factors other than driving scenario and traffic density may affect physiological measures of workload in the driving context. HRV is a relatively consistent and reliable measure of mental workload, but it is very sensitive to respiration, movement frequency, force exertion level, and interactive effects of these factors (Luft et al., 2009). Thus, to make sure that HRV measures reflect the changes in workload, these confounding factors should be excluded or controlled in the experiment.

The effect of work zones on subjective workload was previously researched by Shakouri et al. (2014). In their study, NASA TLX was used to record subjective workload of drivers as they navigated through the CLM and JLM. The results showed that driving through the JLM compared to CLM was 15.3% less demanding.

In this study, unlike physiological measures of workload, the subjective measures were sensitive to the changes in work zone configurations. As expected, participants reported that driving on the road without a work zone is less demanding. Except for physical demand, performance and frustration which were not statistically significant, the other three measures increased with the presence of work zone. The insensitivity of physical demand and frustration can be explained by the short duration of the driving scenarios. Each scenario, on average, took about three minutes, and such short duration may not be enough to induce physical fatigue or trigger frustration. Several studies suggest that frustration and aggressive driving is more influenced by individual differences rather than environmental conditions. In one study, Krahé and Fenske (2002) found that factors such as gender, age, macho personality and driving powerful cars are all significant factors that influence frustration. In another study, Yagil (2001)
found that drivers who are anxious or competitive and highly irritable are more likely to become frustrated and aggressive. In an experiment to study the effect of driving on physical fatigue, Ting et al. (2008) found that excessive driving time is a significant fatigue factor and potential cause of increased physical demand. Similarly, Jagannath and Balasubramanian (2014) found that there is a linear relationship between length of driving time and EMG signals which were recorded from back and shoulder muscle groups. Other researchers associate physical workload to factors such as sleepiness, road surface irregularities, low density traffic, time of day, rain, fog, etc. (Kecklund & Åkerstedt, 1993; McCartt et al., 2000)

The analysis of pairwise comparison of NASA TLX subcomponents revealed that mental demand contributed the most to the total workload. This means that participants perceived the driving task to be more mentally demanding rather than other components of NASA TLX. According to Endsley (1995) in a driving task drivers have to perceive, identify and correctly interpret the relevant objects and elements in the current traffic situation. Drivers then construct and maintain a mental representation of the current situation which forms the basis of driver’s decisions and actions (Endsley, 1995). This process consumes attentional resources and as a result makes driving a mentally demanding task.

Lower HRV is an indicator of higher workload (Kamath & Fallen, 1993). Studies have shown that HF reduces during heavy exertions and awkward postures (Vieira et al., 2012). Since univariate tests found no significant differences between measures of HRV in high and low traffic density, it is difficult to give a conclusive remark on the effect of traffic density on physiological measures of workload. The average LF and HF in high and low traffic density were the same. However, with respect to LF/HF ratio, results showed that driving in low traffic density is slightly more mentally demanding, though this difference was not statistically
significant. In one study, Brookings et al. (1996) examined the effect of air traffic on physiological measures of workload in a group of air force traffic controllers. Similar results were found as eye blink and respiration rate, which were used as physiological measures of workload, were not affected by the different levels of air traffic volume.

The current results regarding the effect of traffic density on subjective workload corroborate previous findings. Shinar (1998) showed that driving in congested roads led to higher frustration and more aggressive driving behavior. Schiessl (2008) found that, due to the fact that drivers in high traffic density are being restricted in the actions available to them, mental load is higher in high traffic flow. Teh et al. (2014) found a linear upward trend in driver workload with increasing traffic flow. Their results revealed significantly higher mental demand, physical demand, time pressure, poorer self-rated performance, greater effort and frustration in medium and high traffic complexity, compared to low traffic complexity. Similarly, in this study, the average subjective mental demand in high traffic density was 22.3% more than that of in the low traffic density. Moreover, as expected, participants found driving in high traffic density more frustrating and more effortful. This suggests that subjective workload, compared to physiological measures of workload, is more sensitive to traffic conditions. In fact, except for performance the rest of NASA TLX subcomponents were all significantly influenced by the levels of traffic density. The subjective workload results indicated that as the driving task required more attention e.g. driving in heavy traffic while paying attention to headway distance, the perception of task demand in drivers increased.

The annals of transportation research are replete with studies that focused on the effect of different work zone configurations on performance variables. In this study, the existence of work zones affected steer and speed variability. In fact, the average speed variability in the CLM was
57% higher than the control group. These sudden changes in speed can be attributed to the causes of accidents in the working zones (Morgan et al., 2010; Paolo & Sar, 2012). Shakouri et al. (2014) compared the CLM and JLM in terms of mean speed and percent maximum braking. They found that the mean speed in two configurations was not different but percent maximum braking was lower in the JLM. Similarly, in this study, no significant difference was found in speed variability between the two work zone configurations but brake variability in the JLM was 26.3% less compared to that of in CLM. This can be explained by the omission of the right of way in the JLM and use of funnel like transition zone. When both lanes have equal right of the way, drivers should be cautious and adjust their speed with both leading vehicles and those that are in the other lanes. As a result, the speed of vehicles in the JLM can be expected to be more homogenous. This homogeneity of speed obviates drivers from excessive braking to adjust the speed in case of sudden variations in the speed of leading vehicles.

Traffic density had a significant effect on performance variables. With the increase in traffic density, steer variability and brake variability increased. These results were expected, as drivers in high traffic density brake more often. Similarly, since in high traffic density, drivers are more influenced by the behavior of other vehicles and are limited to move with the moving speed of traffic, the variability of speed in high traffic density was expected to be lower.

There is considerable debate on the disassociation between physiological workload, subjective workload and performance. Results of the current study showed that first, there is no strong association between physiological measures of workload with subjective workload; and second, there is no association between physiological measures of workload and performance variables; and third, there is a weak relationship between subjective workload and performance
variables. RMSSD was the only physiological variable that significantly correlated with six out of seven subjective variables (RMSSD was not correlated with performance).

Miyake (2001) attributed the disassociation between physiological measures of workload and subjective workload to the task result. He explains further by giving an example of making a ship model. If a participant was given a very complex and delicate ship model, at the end of completion he or she would feel great about it and would rate his/her performance high. But if the model dropped and broke into pieces just before the completion, the participant would feel frustrated and not satisfied with his/her performance. However, if we could record the participant’s physiological workload during the task, the results would be identical in two scenarios as the performance was the same until the end. Thus, feelings of achievement or conception of one’s performance are important in evaluating workload. However, the correlation between such feelings and the physiological responses during the task may be low (Miyake et al., 2009). Similarly, it may be possible that participants’ feelings of accomplishment in finishing the drive or frustration due to road condition, traffic, etc. affected their subjective workload while their physiological measures remained unchanged. In summary, this study was conducted to test three hypothesis:

- Hypothesis 1:
  
  o $H_0$: There is no difference in drivers’ physiological workload at the presence of work zones in different traffic densities.
  
  o $H_a$: There is at least a difference in drivers’ physiological workload at the presence of work zones in different traffic densities.

MANOVA results showed that there was not enough evidence to reject the null hypothesis and therefore we failed to reject the null hypothesis. Although, compared to the rest
period, the physiological measures changed, physiological measures of workload were not
sensitive to the changes in the work zones. With regard to the effect of traffic density on
physiological measures of workload, results showed that RMSSD was the only measure that
detected the change in the traffic density. RMSSD in high traffic density was 9.6% lower than
that of in low traffic density though not statistically different. However, for the remaining
physiological measures, the results were almost the same and not statistically different.

- Hypothesis 2:
  - $H_0$: There is no difference in drivers’ subjective workload at the presence of work
    zones in different traffic densities.
  - $H_a$: There is a difference in at least one measure of drivers’ subjective workload at
    the presence of work zones in different traffic densities.

Results showed that subjective measures of workload were influenced by work zones and
traffic density. Thus, we can reject the null hypothesis in favor of the alternate hypothesis.
Overall, based on total workload, driving on the CLM and JLM were 17.9% and 12.7% more
demanding than the control scenario, respectively. Results showed that drivers going through the
CLM experienced the highest amount of mental demand. Driving through CLM was as
physically demanding as driving through JLM but they were not significantly more demanding
that the control group. Temporal demand was lowest in the control group, followed by JLM with
10.1% and CLM with 23.5% increase. Participants’ self-reported performance showed no
difference between CLM and JLM. However driving on a road with a work zone required
significantly more effort than driving on a road without a work zone. In terms of frustration,
driving through the CLM was the least frustrating. With respect to the effect of traffic density on
subjective workload, results showed that driving in high traffic density is significantly more demanding than driving in low traffic density.

- **Hypothesis 3:**
  - $H_0$: There is no association between subjective workload, physiological workload and performance variables
  - $H_a$: There is at least an association between subjective workload, physiological workload and performance variables

The Pearson correlation results showed that the relationship between physiological measures of workload, subjective workload and performance variables were weak. Only RMSSD was significantly correlated with most of subjective workload variables. This can be explained by a nonlinear relationship between other physiological measures of workload with subjective and performance variables.
6 CONCLUSION

Literature on merge configuration suggests that there are a myriad of metrics to evaluate merge efficiency. Although one configuration may prove to be efficient in one metric, it may perform weakly in another. Thus, knowing which metrics can truly demonstrate the true performance of a merge configuration can assist designers to design safer work zones. In this study the effect of merge configuration and traffic density on three common metrics of physiological measures of workload, subjective workload and performance variables that have been used separately by many researchers were studied together.

Although physiological measures of workload were not affected by the presence of work zone and in different traffic densities, subjective workload and performance variables were influenced. Analysis of relationship between HRV measures of workload and subjective variables suggested some significant but weak correlations between these two groups. There were also some significant but weak correlations between performance variables with subjective workload and physiological measures of workload.

6.1 Limitations

There were several sources of potential limitations in this study. Firstly, the MANOVA had sufficient power but the study lacked sufficient power for the univariate tests to detect any significant effects for physiological measures of workload, even if they exist in reality. For future studies, the priori power analyses to calculate the sample size should be conducted for the univariate tests. Increasing the sample size may improve the power of univariate tests.

Secondly, in this study polar RS800, which was originally designed for cross training, running etc., was used as a hear rate monitoring medium. This device is not specifically designed
for general physical and mental tasks. Although various studies have validated the accuracy of
Polar heart rate monitoring devices (Gamelin et al., 2006; Porto & Junqueira, 2009), we found
that the recorded data gets distorted when the connection between the sensors and watches is
lost. In this study, to avoid forgetting to record the data, participants did not wear the heart rate
monitoring watch and it was the experimenter who started and stopped the recordings. However,
due to distance between experimenter and participants, there were moments that the connection
between the watch and sensor was lost and this represented itself as high RR values (few
outliers) in the recorded data. Thus, it is very crucial to screen the data before doing any analysis
as these extreme values may confound the findings. For future studies, it is recommended to
measure heart rate variability by using portable electrocardiogram (EKG or ECG) which is
specifically designed to measure electrical activity of heart.

Thirdly, as several researchers suggested, individual differences have direct impact on
HRV (Ting et al., 2008; Yagil, 2001). Although in this study information regarding gender and
age were recorded for each participant, they were not included in the hypothesis of this study and
therefore were not considered in the analysis. The reason for the exclusion of these variables was
that, based on priori power analysis, the number of actual subjects (30) was much less than
required number of participants (403) for ANCOVA test with the power of 0.8. Gender, age and
levels of physical fitness must also have affected the stress ratings as individuals differed from
each other. Future research can be done to measure the extent in which individual differences
affect subjective and physiological workload.

Fourthly, the HR at rest was measured while participants were filling the questionnaire.
Reading and answering questions consumes resources and may impose unwanted mental demand
on participants, thus it is recommended that for future studies, heart rate at rest should be measured while participants are sitting quietly.

Fifthly, the length of the time for each scenario was short. The short duration of driving in the driving simulator can be directly associated with not significant results for physical demand. In reality, long haul driving increases the risks of fatigue, drowsiness, reduction in situation alertness and consequently increased physical workload. Therefore, longer driving scenarios may result in higher perceived physical demand.

6.2 Future Directions

The current research found that work zone was not a main effect for physiological measures of workload. To extend this research, the effect of individual differences such as age, gender, personality type, etc. on physiological and subjective workload while driving can be studied. The relationship between individual difference and risky driving behavior is well researched but to the knowledge of author there is a dearth of information on how individual differences may affect subjective and physiological measures of workload while driving. Future studies can focus on filling this gap by answering the following research questions: “Are there any associations among individual differences and workload?” and if yes, “What factors influence drivers’ workload significantly?”. Answering these questions can paint a better picture of driving behavior and may reveal how personality traits and individual differences are linked to driving workload near work zones.

6.3 Contributions

The annals of transportation safety research is replete with studies that have focused on the efficiency of different work zone configurations with respect to performance variables. While
many researchers attest to the importance of other factors such as workload, but few people, in the driving context, have studied how work zones may affect physiological and subjective workload. Results of this research showed that other than performance variables, subjective workload may also be a good metric for evaluation of the efficiency of a merge configuration. This means that if a driver considers the workload of a driving task to be excessive, they may behave as though they are overloaded, even though the task might not be objectively demanding. As a result, drivers’ performance may decline or make mistakes which may consequently lead to accidents.
REFERENCES


Leung, S., & Starmer, G. (2005). Gap acceptance and risk-taking by young and mature drivers, both sober and alcohol-intoxicated, in a simulated driving task. *Accident Analysis & Prevention, 37*(6), 1056-1065. doi: [http://dx.doi.org/10.1016/j.aap.2005.06.004](http://dx.doi.org/10.1016/j.aap.2005.06.004)


Paritala, S. A. (2009). *Effects of physical and mental tasks on heart rate variability*. (Master’s degree), Louisiana State University, Baton Rouge, LA.


Rayaprolu, P. (2010). *Operational And Safety Assessment Of Joint And Conventional Lane Merge Configurations For Freeway Work Zones*. (Master of Science), Louisiana State University and Agricultural and Mechanical College.


APPENDICES

1. LSU IRB Approval

ACTION ON PROTOCOL APPROVAL REQUEST

TO: Laura Ikuma  
MIE

FROM: Robert C. Mathews  
Chair, Institutional Review Board

DATE: April 22, 2014

RE: IRB# 3483

TITLE: The effects of Work zone configurations on physiological and subjective workload


Review type: Full ___ Expedited ___ X ___  Review date: 4/22/2014

Risk Factor: Minimal ___ Uncertain ___ Greater Than Minimal ___

Approved ___ X ___ Disapproved ___

Approval Date: 4/22/2014  Approval Expiration Date: 4/21/2015

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 50

LSU Proposal Number (if applicable): ______

Protocol Matches Scope of Work in Grant proposal: (if applicable)

By: Robert C. Mathews, Chairman

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING – Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU’s Assurance of Compliance with DHHS regulations for the protection of human subjects*

2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.

3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.

4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.

5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.

6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.


8. SPECIAL NOTE:

*All investigators and support staff have access to copies of the Belmont Report, LSU’s Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb
2. Informed Consent Form

**Study Title**
The effects of Work zone configurations on physiological and subjective workload

**Performance site**
Louisiana State University. Full sized LSU driving simulator housed in LSU driving simulator lab in the Department of Civil and Environmental Engineering. Location: Room 2225 Patrick F. Taylor Hall.

**Investigators**
Mahmoud Shakouri Hassanabadi, MIE, (225) 436-4666, 3112 Patrick Taylor Hall, mshako1@lsu.edu
Dr. Laura Ikuma, Associate Professor, MIE, (225) 578 5364, 2156 Patrick F. Taylor Hall, likuma@lsu.edu
Briana Saul, MIE, undergraduate research assistant
Sara Wren, MIE, undergraduate research assistant

**Purpose of the Study**
The purpose of this study is to:
1. To determine the effects of work zone configuration and traffic flow levels on driver’s performance and workload
2. To investigate the association between subjective workload and physiological measures of workload

**Subject Inclusion**
Primarily students, both male and female, from Louisiana State University (LSU), ages 18-60 with a valid driving license.

**Exclusion Criteria**
Individuals that have the following conditions:
1. One who does not have a valid driving license.
2. One who is prone to or show motion sickness

**Number of Subjects: 50**

**Study Procedures**
Each participant will first read this consent form and be given a verbal explanation on the procedures involved in this study. If he/she agrees to the terms of participation, they will sign this form which shows their interest and willingness to participate in the project. At any time during the experiment, if more than normal task operating discomfort is encountered, participants can cease the activity. The experiment starts with participants filling informed consent form and demographic information questionnaire. After that they will be asked to put on a heart rate strap which is worn around the chest. This strap sends heart rate signals to a heart rate monitor watch. In case of female participants, a female assistant will help the participants to put on the belt. After that, each participant will run a test drive and they will be given instruction on how to fill NASA
TLX questionnaire. The experiment includes 6 trials which take about 15 minutes. After each trial, participants’ heart rate and workload is measured by using the heart rate monitor watch and NASA TLX questionnaire, respectively.

**Benefits**
There are no direct benefits; but this experiment may provide information that will yield future improvements in the task of designing and planning to move towards an optimum driving behavior. That will in turn reduce congestions, increase speed and capacity of the roads, satisfied drivers who facilitate emergency evacuations etc.

**Risks/Discomforts**
The only risk is the chances of getting motion sickness. The tasks have been designed to fall within the normal job performance for a good driving condition, so the potential physical or mental discomfort is not expected to be any greater than that, after a typical video game. Participants are encouraged to inform the investigators or the co-investigators, if motion sickness is felt.

**Right to Refuse:** At any time during the experiment, participants have the right to not participate or withdraw from the study. There will be no penalties for withdrawal.

**Privacy:**
Other than as set forth above, participant identity will remain confidential unless disclosure is legally compelled.
Results of the study may be published, but no names or identifying information will be included in the publication.

**Financial Information:** No costs are incurred by subjects in this study.

**Removal:** Participants are expected to comply with the investigator’s instructions. If they fail to comply, they will be removed by an investigator from the experiment.

**Signatures:** The study has been discussed with me and all my questions have been answered. I may direct additional questions regarding study specifics to the investigators. If I have questions about participant’s rights or other concerns, I can contact Robert C. Mathews, Chairman, Institutional Review Board, (225) 578-8692. I agree to participate in the study described above and acknowledge the investigator’s obligation to provide me with a signed copy of the consent form.

__________________________                        __________________
Subject Signature                                                                            Date

________________________________________
Print name
3. Demographic Information Questionnaire

**Instruction:** Please fill an appropriate box for each question.

1. Gender □ Male □ Female

2. Age ---------------------

3. How long have you had your driving license? ---------------------

4. What is the type of your car?

□ 4WD □ Small car □ Sedan car

5. What is your driving experience?

□ <1 □ 1-5 □ 5-9 □ ≥10

6. Estimate the number of miles you drive each year ---------------------

7. During the past year (12 months) have you been involved in any accidents?

□ Yes □ No

8. If yes, how many accidents ---------------------

9. During the past year (12 months) have you had any highway violations?

□ Yes □ No

10. If yes, how many violations ---------------------
4. NASA-TLX

Refer to these descriptions as you complete the Workload Rating sheet.

**Mental Demand: Low/High** How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

**Physical Demand: Low/High** How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

**Temporal Demand: Low/High** How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

**Performance: Excellent/Poor** How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

**Effort: Low/High** How hard did you have to work (mentally and physically) to accomplish your level of performance?

**Frustration Level: Low/High** How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?
Date of Experiment

Instructions: select the member of each pair that provided the most significant source of workload variation in these tasks.

### Scale Comparison

<table>
<thead>
<tr>
<th>#</th>
<th>Physical Demand</th>
<th>Mental Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Temporal Demand</td>
<td>Mental Demand</td>
</tr>
<tr>
<td>2</td>
<td>Temporal Demand</td>
<td>Physical Demand</td>
</tr>
<tr>
<td>3</td>
<td>Performance</td>
<td>Physical Demand</td>
</tr>
<tr>
<td>4</td>
<td>Temporal Demand</td>
<td>Frustration</td>
</tr>
<tr>
<td>5</td>
<td>Temporal Demand</td>
<td>Effort</td>
</tr>
<tr>
<td>6</td>
<td>Performance</td>
<td>Mental Demand</td>
</tr>
<tr>
<td>7</td>
<td>Frustration</td>
<td>Mental Demand</td>
</tr>
<tr>
<td>8</td>
<td>Effort</td>
<td>Mental Demand</td>
</tr>
<tr>
<td>9</td>
<td>Frustration</td>
<td>Physical Demand</td>
</tr>
<tr>
<td>10</td>
<td>Effort</td>
<td>Physical Demand</td>
</tr>
<tr>
<td>11</td>
<td>Temporal Demand</td>
<td>Performance</td>
</tr>
<tr>
<td>12</td>
<td>Performance</td>
<td>Frustration</td>
</tr>
<tr>
<td>13</td>
<td>Performance</td>
<td>Effort</td>
</tr>
<tr>
<td>14</td>
<td>Effort</td>
<td>Frustration</td>
</tr>
</tbody>
</table>
Workload Rating

Instructions: Place a vertical mark on each scale that represents the magnitude of each factor in the task you just performed.

**Mental Demand**

Low

High

**Physical Demand**

Low

High

**Temporal Demand**

Low

High

**Performance**

Excellent

Poor

**Effort**

Low

High

**Frustration Level**

Low

High
5. Motion Sickness Questionnaire

Thank you for taking the time to complete this questionnaire. This questionnaire is a part of our Compliance Policies Procedures to Institutional review board (IRB). Your safety during the experiment is our highest concern and the information you provide help us to monitor and control your safety throughout the simulation. Please take your time and answer the question carefully. Should you have any questions please don’t hesitate to contact us at mshako1@lsu.edu or call 225-436-4666.

Directions:

Please read the symptoms provided in the table below and tell us if any of those have. You can show the severity of the symptom by marking the corresponding number. 0 means you don’t have that symptom and as the number goes up the severity increases proportionally.

<table>
<thead>
<tr>
<th>Motion Sickness Assessment Questionnaire (MSAQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Do you feel ....</strong></td>
</tr>
<tr>
<td>Sick to stomach</td>
</tr>
<tr>
<td>Faint-like</td>
</tr>
<tr>
<td>Annoyed/irritated</td>
</tr>
<tr>
<td>Sweaty</td>
</tr>
<tr>
<td>Queasy</td>
</tr>
<tr>
<td>Lightheaded</td>
</tr>
<tr>
<td>Drowsy</td>
</tr>
<tr>
<td>Clammy/cold sweat</td>
</tr>
<tr>
<td>Disoriented</td>
</tr>
<tr>
<td>Tired/fatigued</td>
</tr>
<tr>
<td>Nauseated</td>
</tr>
<tr>
<td>Hot/warm</td>
</tr>
<tr>
<td>Dizzy</td>
</tr>
<tr>
<td>Like I am spinning</td>
</tr>
<tr>
<td>As if I might vomit</td>
</tr>
<tr>
<td>Uneasy</td>
</tr>
</tbody>
</table>
6. Data Screening

6.1 RMSSD before Removing the Outliers
6.2 RMSSD after Removing the Outliers
6.3 LF
6.4 HF
6.5 LF/HF Ratio before Removing the Outliers
6.6 LF/HF Ratio after Removing the Outliers
6.7 Mental Demand
6.8 Physical Demand

[Bar charts showing frequency distribution for Control Group, CLM, and JLM]
6.9 Temporal Demand
6.10 Performance

[Bar charts showing frequency distribution for Control Group, CLM, and JLM.]
6.11 Effort
6.12 Frustration
6.13 Total Workload

The diagram shows the frequency distribution of total workload for three different groups: Control Group, CLM, and JLM. The x-axis represents the range of workload values, while the y-axis represents the frequency. The bars indicate how many instances fall into each workload range for each group.
6.14 Steer Variability (degrees)
6.15 Brake Variability (N)
6.16  Speed Variability (m/s)
VITA

Mahmoud Shakouri Hassanabadi, was born in 1985 in Tehran, Iran. He received his Bachelors’ degree in Architectural Engineering from Islamic Azad University in Iran. After that he worked in a construction company for three years and then moved to Malaysia to do his Master’s degree in Construction Management. He graduated from Universiti Teknologi Malaysia in 2012 and few months later he joined the department of mechanical and industrial engineering at LSU in fall 2012 as a research assistant. During his studies at LSU he held several positions as teaching assistant for Occupational Biomechanics lab (IE 7467), Safety Engineering (IE 4462), Methods of Engineering lab (IE 2400), and Human Factors Engineering lab (IE4761). He is a member of Human Factor and Ergonomics Society and Institute of Industrial Engineers. He is a recipient of Gulf Coast Center for Evacuation Transportation Resiliency Graduate Scholarship.

He has several publications in Accident Analysis & Prevention, The Transportation Research Record and 2013 Industrial and Systems Engineering Research Conference.