Effects of Task Automation on the Mental Workload and Situation Awareness of Operators of Agricultural Semi-Autonomous Vehicles

By

Behzad Bashiri

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Biosystems Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

Degree: Doctor of Philosophy

University: University of Manitoba

Department: Biosystems Engineering

Title: Effects of Task Automation on the Mental Workload and Situation Awareness of Operators of Agricultural Semi-Autonomous Vehicles

Author: Behzad Bashiri

Supervisor: Dr. Danny D. Mann

ABSTRACT

The effects of in-vehicle automation and driving assistant systems on the mental workload and situation awareness of drivers have been the interest of many studies; some of the implications of automation in such man-machine systems have been identified. Due to the introduction of advanced automated systems in agricultural machinery, farmers are currently working with semi-autonomous vehicles. A human factors perspective on the design of these systems will ensure safe and efficient operation of such man-machine systems.

In this study, a systematic approach was utilized to address human factors issues associated with operating a semi-autonomous agricultural vehicle, and to provide design recommendations. The study was carried out in three stages. First, a task analysis was used to identify tasks associated with operating an agricultural vehicle and to select appropriate experimental variables. Next, a preliminary experiment was performed to validate the test procedure and measurement techniques. Finally, the main experiment was administered. Experiments were conducted using the Tractor Driving Simulator located in the Agricultural Ergonomics Laboratory at the University of Manitoba. Thirty young experienced tractor drivers participated in this study. The experiment investigated the effects of i) vehicle steering task automation (VSTA) and ii) implement control and monitoring task automation (ICMTA) on mental workload and situation awareness of drivers.

It was found that ICMTA significantly affected situation awareness (and its underlying components) of the operator. The situation awareness of drivers increased as the automation support level increased, but the highest level of automation, where the

participants were out of the task loop, resulted in low situation awareness, similar to the condition with no automation support. VSTA only reduced the attentional demand of the situation, one of the three components of the situation awareness, which had negative effect on overall situation awareness.

Based on the results from a subjective mental workload measure, moderate levels of mental workload were reported when the participants were involved in the implement control and monitoring task loop. The highest level of ICMTA reduced the average mental workload by 18%. Reaction time of drivers and number of errors committed by drivers both decreased as the automation level increased.

AKNOWLEDGMENT

First and foremost, I am grateful to God for the good health and wellbeing that were necessary to complete this dissertation. I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family.

I wish to thank my committee members who were more than generous with their expertise and precious time. I would like to express my deepest gratitude to my advisor, Dr. Danny Mann, for his excellent guidance, care and patience, providing me with an excellent atmosphere for doing research. I would also like to extend my sincerest thanks to Dr. Jason Morrison and Dr. Tariq ElMekkawy for agreeing to serve on my committee. I would also like to thank my external examiner, Dr. Greg Jamieson for his insightful questions.

Special thanks go to the people who volunteered to participate in my study, and the staff and technicians in the Department of Biosystems Engineering for their continued support. I am particularly thankful to my fellow researchers in the Agricultural Ergonomics Lab, Aadesh Rakhra and Davood Karimi.

Many people were indirectly involved in my research, and their support and contributions have also been essential for this dissertation. I would like to thank Dr. Garry Crow and Dr. Hamidreza Khazanei from the Department of Animal Science for the helpful discussions on experimental design and analysis. I would also like to thank Dr. Annie Pauzie for the support and advice on DALI technique, and Dr. Michelle Porter from the Faculty of Kinesiology and Recreation Management for loaning the heart rate monitor used in this study.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant for the first 3 years and partially for years 4 and 5. In addition, I was honored to be a University of Manitoba Graduate Fellow for my 4th year.

My time at University of Manitoba was made enjoyable in large part due to the many friends and groups that became a part of my life. I am grateful for time spent with these people.

Finally, I would like to thank my family for all their love, support, and encouragement; to my parents who raised me with a love of science and supported me in all my pursuits and to my brothers and sisters for all of the emotional support and caring that they have provided me.

DEDICATION

I dedicate this work to my parents who instilled the love of knowledge in my brothers and sisters as well as me. To my mother, who passed away very suddenly during the course of this work; you will always be remembered for your selfless heart and rich generosity. I will be forever grateful to you for sacrificing your life to give of yourself to your children. To my father; you remain a hardworking man who believes in education, always wanting me to get the highest academic degree and trusting in my ability and intelligence to do so. Thank you both for believing in my education and for encouraging my affinity for science.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PUBLICATIONS FROM THIS STUDY

Some results of this study have been published in scientific journals or conference proceedings. A list of these publications as well as papers presented in conferences, in the form of oral or poster presentation, is provided below. Permissions, if needed, were obtained from the respected publisher for reusing materials in this thesis.

1- Bashiri, B., D.D. Mann. 2015. Impact of automation on drivers' performance in agricultural semi-autonomous vehicles. *Journal of Agricultural Safety and Health*, 21(2): 129-139.

2- Bashiri, B., D.D. Mann. 2014. Automation and the situation awareness of drivers in agricultural semi-autonomous vehicles. *Biosystems Engineering*, 124: 8-15.

3- Bashiri, B., D.D. Mann. 2014. Heart rate variability in response to task automation in agricultural semi-autonomous vehicles. *The Ergonomics Open Journal*, 7: 6-12.

4- Bashiri, B., D.D. Mann. 2013. Drivers' Mental Workload in Agricultural Semi-Autonomous Vehicles, In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57: 1795-1799.

5- Bashiri, B., D.D. Mann, A. K. Rakhra. 2013. Automation and drivers' performance in agricultural semi-autonomous vehicles. *CSBE Conference for Interdisciplinary Engineering*, July 7-10, Saskatoon, SK, Canada.

6- Bashiri, B., D.D. Mann, D. Karimi. 2011. Function oriented task analysis of agricultural vehicles. *Annual Meeting and Technical Conference of Human Factors and Ergonomics Society (Europe Chapter)*, October 19-21, Leeds, England.

7- Bashiri, B., D. Karimi, D.D. Mann. 2011. Driving Simulation for Tractor-Air seeder system. *2011 CSBE/SCGAB Annual General Meeting and Technical Conference*, 10-13 July 2011, Winnipeg, MB, Canada.

8- Bashiri, B., D.D. Mann, D. Karimi. 2011. Hierarchical Task Analysis of Driving a Tractor Air Seeder System (TAS): Determining TAS Simulator Requirements. *2011 CSBE/SCGAB Annual General Meeting and Technical Conference*, 10-13 July 2011, Winnipeg, MB, Canada.

ABBREVIATIONS

Acq	Information acquisition
Act	Action implementation
Ana	Information analysis
D	Demand on attentional resources
DALI	Driving Activity Load Index
Dec	Decision and action selection
ICMT	Implement control and monitoring task
ICMTA	Implement control and monitoring task automation
LOA	Level of automation
Man	Manual
S	Supply of attentional resources
SA	Situation Awareness
SART	Situation Awareness Rating Technique
TAS	Tractor air-seeder system
TDS	Tractor Driving Simulator
TLX	Task load index
U	Understanding of the situation
VSTA	Vehicle steering task automation

Chapter 1

GENERAL INTRODUCTION

1.1 Introduction

"What happens when you fall asleep with Autotrac?" was the headline for an accident reported in June 2007 (Fone 2007). The operator of an agricultural tractor and an air seeder, equipped with a GPS auto-steer system, fell asleep while working on a field. The machines ended up in a tangle with a high-voltage electrical pylon. The number of similar incidents is increasing, possibly because new automation systems reduce the role of human operators in the operation of agricultural semi-autonomous vehicles.

Operation of an agricultural vehicle on a field is a continuous task that requires much physical and mental effort. It usually involves two co-primary tasks of i) driving the agricultural vehicle and ii) monitoring and controlling the implement being powered by the agricultural vehicle. A variety of automated systems have been introduced in these vehicles to enhance operation performance and reduce the operator's workload. Automatic steering systems, for instance, allow drivers to delegate the steering task while driving on a straight path toward a headland, allowing them to assign more attention to the implement control and monitoring task (ICMT). The ICMT also has been the subject of automation and its sub-tasks have been automated - partially or entirely. Although physical workload is reduced with these automated systems, the mental workload and situation awareness of the driver, when dealing with such systems, remains unknown. Studies in different domains have proven that the design of any automated components in man-machine systems may cause human errors if the human factors perspective is ignored. Considering the huge market of agricultural vehicles in Canada and the number of operators in this industry, the significance of designing agricultural vehicles from a human factors perspective becomes more clear.

In presenting a system for agricultural tractor automation, Stentz et al. (2002) raised the following questions that address automation at a system level for a system of people and computer-controlled machines working together:

i. How should operations be divided between the human and the machine?

ii. What is the best way for the human to interact with the machine?

iii. How can the machine be made productive and reliable for the tasks it is given?

As the technology is advancing, a growing trend of studies can be found on addressing the third question. The present study is an attempt toward answering the first two questions in agricultural practice. It focussed on the variation of mental workload and situation awareness of operators when working with agricultural semi-autonomous vehicles. The effects of in-vehicle automation are investigated systematically using human factors theories. Efficient allocation of functions between operators and semi-autonomous agricultural vehicles is envisioned as the ultimate goal of this work.

1.2 Thesis overview

The following chapter reviews relevant literature pertinent to the research questions. It provides the necessary background information for the study. Chapter 3 presents the proposed objectives as well as the framework of the study that was adapted from the model for types and levels of automation suggested by Parasuraman et al. (2000).

The first stage of the framework is discussed in chapter 4; this chapter presents the task analysis procedure used for better understanding the tasks associated with the operation of an agricultural vehicle. The findings from this stage were used to select the independent variables of the study and to inform the experimental design. Chapter 5 describes the experimental design and experimental procedures. Information regarding the experiment participants and the research tools are also provided. Results from the preliminary investigation and the subsequent full set of experiments are provided in Chapters 6 and 7, respectively. Correlations between mental workload, situation awareness, physiological response and performance of the drivers are also discussed in these chapters. In Chapter 8, a discussion of the results is provided and results are compared with findings from relevant studies. Chapter 9 summarizes the outcomes of the entire research project and describes limitations of the study. Finally, Chapter 10 lists some recommendations for future studies.

Chapter 2

LITERATURE REVIEW AND THEORY

2.1 Introduction

This chapter provides a general overview of previous research on human factors in human-machine systems. It introduces human factors issues in working with automated systems. The goal is to use this knowledge for identifying potential human factors issues in agricultural semi-autonomous vehicles and subsequently identifying how human factors tools can be used for investigating these issues in automation design of such vehicles. Differences between on-road driving tasks and on-field operation have been explained to vindicate the necessity for this study. In addition, some of the automated systems that have been implemented in agricultural vehicles, and how they have changed the traditional driving task of such vehicles, have been outlined. In the literature review, a framework for the case study has been explained that comprises the focus of the research described in this thesis. Finally, a review of simulator use for automation studies has been included to set the stage for the use of a tractor driving simulator in this study.

2.2 Automation and human errors

We may be tempted to think that automation means that the machine will replace the human because it is better able to complete the task. A more appropriate view, however, is that the system should be designed so that the human and the machine are able to work together and to complement one another (Hollnagel & Bye 2000). Researchers have provided lists of automation levels in which a fully automated system was defined as a system where the operator is completely out of the control loop; the minimal level of automation was described as a system where the automation only presents basic data filtering or recommendations for the human to consider (Parasuraman et al. 2000; Riley 1989; Sheridan & Verplank 1978). Norman (1990) stated that the current level of automation in industry does not provide adequate feedback to the human operator, so automation can cause human error when the situations exceed the capabilities of the automatic equipment. When working with automated systems, operators may experience loss of situational awareness (Cummings 2004; Endsley 1999; Stanton & Young 2005; Walker et al. 2008), vigilance decrement (Finomore et al. 2009; Parasuraman 1986), complacency (Kaber & Endsley 2004), and skill degradation (Billings 1996; Wickens & Hollands 2000); any of these situations can lead to human errors in man-machine systems.

Situation awareness is "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley 1996). A high level of situation awareness enables the human to act effectively and timely, even with very complex and challenging tasks (Endsley 2013). A study by Walker et al. (2008) revealed a direct relation between feedback and situation awareness of the driver. By conducting experiments in simulated and naturalistic driving conditions, they found that current trends in vehicle design might show a generalized trend towards decreasing situation awareness of the driver.

Vigilance decrement is an effect of repetitive observation during boring monitoring tasks and refers to a decrease in reaction times or an increase in error rates (Pattyn et al. 2008). It typically happens after 20–30 min of continuous work, however, under certain conditions it can occur in as little as 5 min (Caggiano & Parasuraman 2004). Two opposing theories regarding underlying causes of decrement in vigilance, according to 50 years of research, are i) withdrawal of the supervisory attentional system due to underload (insufficient workload) and ii) limited information processing resources due to mental fatigue or over-load (Helton & Russell 2011). A possible solution for alleviating such an effect is the use of an automated system which can detect the situation and take over the function in response to situational demands and operator performance (Freeman et al. 2004).

Complacency, or over-reliance on automation, happens as operators rely on highly reliable automation, but still there is a possibility of automatic system failure without warning; in other words, complacency denotes the development of a false sense of security (Billings 1991). Bahner et al. (2008) conducted a laboratory experiment using a process control simulation. They provided the people who participated in the research with an automated decision aid for fault diagnosis and management. They measured complacency directly by measuring the participants' information sampling behavior (i.e., the amount of information sampled in order to validate the automated recommendations). They found that teaching automation failures during training significantly reduces the complacency of the operator. Moray (2003) commented that a "failure" to monitor is more likely to be a eutectic strategy than complacency, so the problem is the system design rather than human fault and even optimal sampling cannot identify all abnormal events. This view suggests a radical system re-design in order to guarantee the desired level of performance.

Undoubtedly, automation of activities will result in decay in the skill of the operator due to lack of practice (Wickens et al. 1997). This effect becomes more substantial as the reliability of the automation increases (i.e., as the opportunity for the operator to practice manual control is minimized) (Dekker 2004). This will be very important when the automation reaches its limits and the operator is required to assume control of the task. According to Balfe et al. (2011), a key principle for the design of cooperative automated systems is that the automation should incorporate a method to guard against operator skill degradation. Some researchers believe that, although automation may result in skill degradation, it gives new skills to the operator (Form 1987). On-screen monitoring skill, for instance, is one of the new skills that an operator can gain over the traditional monitoring task.

2.3 Automation in agricultural vehicles

In recent years, many automated systems have been introduced in agricultural vehicles to increase their productivity (Edan et al. 2009). It is difficult to generalize the application of automation in agricultural vehicles due to their wide variety and diversity throughout the world. As an example, for a modern tractor-air seeder system, we can list numerous automatic features such as automatic steering, control of working depth, control of travel speed, control of seed and fertilizer application rate, and empty tank warning system. Automated vehicle navigation systems represent perhaps the most important addition to agricultural vehicles because this form of automation has made it possible to have a precise and economic operation by offering features such as steering assistance, automatic steering, and route planning. These automated systems change the nature of the driving task in such vehicles. Considering the aforementioned negative effects that automated systems may cause on human operators, as Lang et al. (2009)

states, human factors must be considered to ensure the safe and efficient operability of these machines.

2.4 Driving tasks and automation

Driving tasks are one of the important subjects of automation. Lunenfeld (1989) defined the driving task as an information-decision action in which real-time information received in-transit combined with prior information and knowledge are used to make decisions and perform actions in a continuous feedback process. Traditionally, the driving task, as stated in Lunenfeld (1989), was subdivided into three discrete, interrelated subtasks: control, guidance, and navigation. Vehicle control includes controlling driving variables such as vehicle speed and position. Guidance contains road-following and safe path maintenance subtasks. Navigation is a two-phase task that consists of pre-trip and in-transit route finding. In an extensive study of the influence of advanced traveler systems on driving tasks, Wheeler et al. (1996) added the two additional driving subtasks of i) vehicle system operation and monitoring, and ii) reacting to emergencies. Monitoring the information provided inside the vehicle regarding the vehicle condition or roadway condition is a visual task that must not be ignored when analyzing driving tasks. At times, drivers need to react to emergencies. Even though this is not a routine activity and does not happen at predicted times, it must still be counted as a subtask of the driver. Each of these subtasks involves some physical or mental activities that the driver or automated vehicle should perform.

With advances in technology, some traditional driving activities are now being changed. Features such as automatic transmission and automatic steering, for instance, have replaced manual gear change and manual steering. Navigation assistant systems have added some monitoring tasks to the driving tasks. Generally, the benefits of automation in vehicles can be listed as comfort, economy and safety (Hahn 1996). Parking assist system, cruise control, power locks, remote access, power windows, and power trunk are some of the comforts introduced in vehicles through automation. Automatic transmission reduces fuel consumption. Similarly, a navigation system can reduce fuel consumption by showing the fastest and most economical route to a destination. Safety is always a main concern and the subject of in-vehicle automation. A congestion assistant system, for instance, ensures safe driving in heavy traffic situations.

As driving tasks are automated, both benefits and problems are introduced to the human operator (Stanton & Marsden 1996). Workload of the driver is one of these problems; however, there is some controversy about the effects of automation on workload. Some researchers believe that automation decreases physical workload, but increases mental workload due to increased attention demand. Research results by Stanton et al. (2001) suggested that despite the nature of the workload changing (i.e., from physical workload to mental workload), the overall workload for driving an automated vehicle is the same as for manual driving. Other researchers have an opposite view; automation may reduce mental workload creating an underload which poses a different problem (Young & Stanton 1997). According to Parasuraman et al. (1993), underload causes boredom and accordingly some critical monitoring functions may be performed inappropriately. When an automation failure scenario happens, a significant proportion of drivers cannot effectively resume control of the vehicle effectively (Young & Stanton 1997). The advantages and disadvantages of in-vehicle automation (i.e., advanced cruise control) on performance, workload and attention allocation of the driver have been demonstrated (Ma & Kaber 2005).

2.5 Driving tasks of agricultural vehicles

Although there are many human factors studies on driving tasks of on-road vehicles, research studies focused on off-road vehicles, particularly agricultural vehicles, are required. Agricultural vehicles have been designed to enable various agricultural operations, and are often required to function on unprepared and changing terrain at relatively high speed (Q. Zhang et al. 1999). These unique driving conditions for agricultural vehicles have resulted in different forms of automation. For instance, a navigation guidance system supports both on-road and on-field operations, to guide the vehicle both when traveling on roads and during field operation (Q. Zhang et al. 1999). Considering the importance of agricultural products for human beings, the necessity of studying these driving tasks is clear. Unfortunately, there are only a few studies on the analysis of driving tasks for agricultural vehicles to this point.

In one study, task analysis was used to design a risk mitigation system for reducing rollover accidents of farm tractors while an implement is connected to them to move a load (Etzler et al. 2008). Performing this task analysis led to a solution that provides force feedback to the operator in risky situations. Following this study, Marzani et al. (2009) used a methodology based on hierarchical task analysis and function allocation techniques to describe the requirements of an agricultural or off-highway human-machine system that is able to recognize potential risky situations and consequently prevent them. Their methodology was composed of three steps: i) establishing critical tasks and then analyzing them through a hierarchical task analysis technique, ii) identifying the subtasks suitable for partial or total automation, and iii) generating and evaluating different alternatives for automation (function allocation) based on the York function allocation method discussed in Dearden et al. (2000). The York method, which is a dynamic function allocation (see section 2.5.2.2) method, uses scenarios as basic units of function allocation to make allocation decisions. The method first makes some mandatory allocations for machine and operator and then considers scenarios to allocate resources (human or machine) to the functions. After this stage, candidates for dynamic function allocation are identified. A possible limitation of the York method is the rationale for choosing either mandatory and partial automation function allocation; the decision making process is highly dependent on the preference of the decision maker and is not a generalized regulation.

Lang et al. (2009) carried out a pilot study at the Technical University of Braunschweig, Braunschweig, Germany, to provide an initial evaluation of relevant human factors in agricultural machinery. As a case study, they chose a self-propelled forage harvester. They identified human factors problem areas associated with operating this agricultural vehicle. Unfortunately, there is no detailed result published from their study.

In another study, Dey and Mann (2010) performed a task analysis to measure the workload of the operator of an agricultural sprayer that is equipped with a lightbar navigation device and a sprayer equipped with an auto-steer navigation device. The heart rate variability measurement showed higher mental workload associated with operation using the lightbar navigation device, however, eye-glance behavior showed that operators of the auto-steer navigation device spent more time viewing the lightbar than operators of the lightbar navigation device did. In the absence of any further research, there is a need for a comprehensive study of the operator's tasks when using semi-automated agricultural vehicles to obtain an appropriate function allocation.

Although we may find some analogies among driving tasks of on-road vehicles and driving tasks of semi-autonomous agricultural vehicles, unique diffrences exist between them when considering the use of agricultural vehicles in the field setting. Unlike the on-road driving context, there is usually only one vehicle working on a field (i.e., no other traffic) and field scenery remains relatively constant because work is confined to the boundary of the field. Driving speed varies according to the type of operation, but rarely exceeds 10 km/h; in-field driving can be considered very slow compared to on-road driving. Consequently, driving an agricultural vehicle includes two primary tasks: i) guiding the agricultural machine across the field and ii) monitoring and controlling the functioning of the agricultural machine. Drivers should distribute their attention between these tasks. Given that automation has diverse and sometimes contradictory effects on driving behaviour in on-road vehicles, and differences exist between on-road driving and on-field operation of agricultural vehicles, specific human factors research is required for agricultural vehicles to identify effects of automation.

2.6 Man-machine interaction and human factors studies

In 1951, Fitts proposed lists of tasks that are most appropriately completed by humans and tasks that are most appropriately completed by machines. Humans might be perceived to have an advantage at detecting small amounts of visual, auditory, or chemical energy; perceiving patterns of light or sound; and both improvising and exercising judgment. Machines will typically be able to respond to control signals more quickly and more precisely. Sheridan (2002) proposed that “the human should be left to deal with the big picture while the computer copes with the details.” There have been similar statements regarding the assignment of functions between humans and machines. The underlying premise of these proposals is that attaining optimal system performance requires assigning of functions to each agent (i.e., human or machine) based on their skills and capabilities (Sanchez et al. 2010). It has been stated that the first step in deciding an appropriate allocation of functions between a human operator and a machine is to determine the goals to be achieved and the tasks necessary to achieve those goals (Hollnagel & Bye 2000). In fact, task analysis is the basis of deciding allocation of functions.

2.6.1 Task analysis

Task analysis is a methodology that covers a range of techniques that help the analyst to obtain descriptions of tasks, to organize and represent the tasks, and then to evaluate systems against functional requirements (Crystal & Ellington 2004). This thesis uses the meanings of task and analysis provided by Kieras (1997). Task refers to the operator’s job or work activity - what he/she is attempting to accomplish. Analysis refers to a relatively systematic approach to understanding the operator’s task that oversteps unaided intuitions or speculations, and tries to document and describe exactly what the task includes.

By conducting task analysis, we can integrate the human element into system design and operation more effectively and efficiently in terms of system safety, productivity and availability issues (Kirwan & Ainsworth 1992). We can start the analysis by providing a description of any particular activity or function related to the analysis objective. Some of the information and data collection methods that have been found to produce useful information about tasks are: observation of user behavior, study of critical incidents and major episodes, questionnaires, structured interviews, and interface surveys (Kieras 1997). After obtaining the required information about tasks, asking the "how" question causes the analyst to decompose a function into smaller elements or sub-tasks; asking the "why" question compels the analyst to identify any higher-level activities that need to be considered.

Kirwan and Ainsworth (1992) named six major human factors issues for which task analysis can be used to assess and assist their adequacy to ensure system success. These issues are: a) allocation of function, b) person specification, c) staffing and job organization, d) task and interface design, e) skill and knowledge acquisition, and f) performance assurance. A list of task analysis techniques for each of these issues can be found in Kirwan and Ainsworth (1992).

2.6.1.1 Task analysis approaches and techniques

Today there are many task analysis techniques available (Lafreniere 1996; Limbourg et al. 2001; Richardson et al. 1998). Kirwan and Ainsworth (1992) have provided 25 methods and case studies in their book. Overall, there are three approaches to task analysis: a technical (ergonomic) approach, a conceptual (information processing) approach, and a contextual (work process) approach (Crystal & Ellington 2004).

2.6.1.1.1 Technical approach

For modeling the ergonomic aspect, hierarchical task analysis can be used (Crystal & Ellington 2004; Stanton & Middlesex 2006). Hierarchical task analysis, a method developed by Annett and Duncan (1967), involves identifying the overall function in a top down fashion of the task, sub-tasks and the conditions under which they should be carried out to achieve that function (Shepherd 1998). According to Kirwan and Ainsworth (1992), for human factors issues such as function allocation and interface design, hierarchical task analysis is an ideal method to represent task activities and to record task knowledge. Hierarchical task analysis has been widely used for interface design and evaluation, for allocation of function, for assessment of workload, for job aid design, and for error prediction (Stanton & Middlesex 2006). Shepherd (1998) demonstrated that this method is also beneficial for analysis of cognitive tasks. Hodgkinson and Crawshaw (1985) applied this method to the task of mixing sound to identify the human factors and ergonomics requirements of design and evaluation of sound-mixing consoles. In the automotive domain, Wheeler et al. (1996) used a function-oriented hierarchical task analysis method to develop detailed human factors design guidelines for Advanced Traveler Information Systems and Commercial Vehicle Operations. In the agricultural domain, Marzani et al. (2009) used a methodology including hierarchical task analysis to describe the requirements for a human-machine system able to recognize potential risky situations and prevent them.

2.6.1.1.2 Conceptual approach

Cognitive techniques are used to model the information-processing (conceptual) tasks (Crystal & Ellington 2004; Gordon & Gill 1994). A cognitive task is a set of related mental activities, which are unobservable, used to achieve a goal (Wei & Salvendy 2004). According to the Schraagen et al. (2000), cognitive task analysis is defined as ‘the extension of traditional task analysis techniques to yield information about the knowledge, thought processes, and goal structures that underlie observable task performance'. This kind of task analysis is hard to do, is both time consuming and labor intensive, and therefore, expensive (Redding 1990; Wei & Salvendy 2004). Cognitive task analysis has been used in intelligent tutoring system development, decision support system design, and knowledge elicitation and acquisition for expert systems (Ryder & Redding 1993). Hamilton and Clarke (2005) developed a model of train driver information processing, utilising cognitive task analysis and modeling techniques, to understand and manage the driver's interaction with the infrastructure through line-side reminder appliances.

2.6.1.1.3 Contextual approaches

Activity theory is used to model the contextual approaches to task (Bedny & Harris 2008; Crystal & Ellington 2004). In activity theory, the activity is the basic unit of analysis (Spinuzzi 1997); any task can be broken down into actions, which are further subdivided into operations. This can provide an understanding of the steps that the user requires to perform a task (Kuutti 1996). The consequence of analysis using activity instead of task is that the scope and complexity of analysis is greater. This theory accounts for learning effects and extends the scope of technology, but requires a high level of abstraction and is difficult to apply systematically (Crystal & Ellington 2004).

2.6.2 Function allocation

Function allocation in human-machine systems is defined as dedicating function between the human operator and the automated system in such a way that we reach the goal more efficiently and safely. For allocation of function, effectual principles and methods are needed since the level of automation has indispensable consequences for the human–system interaction and workload (Bye et al. 1999). In other words, when designing work systems, determining the degree of automation that includes the allocation of functions between human operator and automated system, is a key point (Grote et al. 1995). Since the origins in the early 1950s, the use of function allocation in human factors and automation studies is entering its' fourth generation (Lagu & Landry 2011). Generally, the generations or theories of function allocation can be classified into two types: static function allocation and dynamic function allocation. Static function allocation exists when allocation of the functions between the human and the machine is performed only once and remains invariant until all functions are completed. On the other hand, dynamic function allocation gives flexibility when assigning the functions.

2.6.2.1 Static function allocation

The first generation of function allocation theories, from the early-1950s to the mid-1980s (Lagu & Landry 2011), are referred to as being 'static'. Static function allocation only considers the allocation of a function to the machine or to the man according to ability as suggested by Fitts (1951), until all functions have been assigned (Scallen & Hancock 2001). The Fitts list included a list of statements about whether a human or a machine performs a certain function better. Thus, this generation of function allocation can be based on left-over criteria (Bye et al. 1999), where the main principle is that only the functions that have not been automated or cannot be automated are to be assigned to the operator (Grote et al. 1995; Inagaki 2003).

This method of function allocation showed some limitations in human-machine studies; it does not consider changes in the type of role and workload of the human operator or evolution of the automated agent. At the time, the lack of computing capabilities caused only highly repetitive and simple tasks be assigned to the automatic machine. Today, automatic systems can be employed to perform even decision-making tasks. With this advancement in technology, static function allocation also shows further limitations; the effect of over-reliance on automation or complacency is abandoned (Lagu & Landry 2011). An automatic system can provide suggestions to the operator, but over-reliance on the automated system by the operator may cause the operator to ignore a problem – another possible system failure. Reviews by Hancock and Scallen (1998) and Dearden et al. (2000) further discuss these sources of possible failure. Despite the extensive criticism that the Fitts list has received, de Winter and Dodou (2014) explained the reasons that the list is still such a pervasive factor in function allocation research. Although the Fitts list perhaps is no longer completely valid because of technological improvements, they showed that the Fitts list fulfils six important criteria for appraising scientific theories: plausibility, explanatory adequacy, interpretability, simplicity, descriptive adequacy, and generalisability.

2.6.2.2 Dynamic function allocation

In the early 1980s, the first generation of dynamic function allocation (i.e., the second generation of function allocation) was introduced (Lagu & Landry 2011). The idea behind the second generation of function allocation is that the total workload in a system is constant, so it is possible to reallocate functions according to the situation of the system agents in order to maximize system performance (Kaber & Endsley 2004). In this scenario, the automated system is responsible for realizing the human's situation. When the human is experiencing underload or the machine has failed, the control of some functions should be moved to the human agent. When the human agent is overloaded, the machine should assume control of some functions from the human (Lagu & Landry

2011). An automated agent monitors human workload changes at fixed time intervals and, if there is a need to alter the system status, the automated agent implements the change (Lagu & Landry 2011). This means that the level of automation is continuously being changed in response to changing conditions (Inagaki 2003) – for this reason, dynamic function allocation is also known as adaptive automation.

The third generation of function allocation fixes the problems associated with fixed time intervals for checking system status. In this generation of function allocation, a real-time mechanism is used to set the function reallocation. Similar to the previous generation, the automated agent is in charge of detecting the necessity for changing the level of automation. For this function allocation, there are four major categories of motives for reallocation. The first category is 'critical events' which refers to events that seek some user interaction. Reallocations are based on a measure of mental state (Scerbo 1996). The second category is 'operator physiological assessment' which refers to real-time assessment of human physiological factors such as heart rate variability, stress variation, and eye dilation, which can be used to make reallocation decisions (de Brunélis et al. 2008). The third category is 'system performance model' which covers the second generation of function allocation. Prior knowledge of how the workload will vary is used by the system developer to make reallocation decisions. The fourth category is 'hybrid methods' which refers to strategies that are a combination of two or more of the previous categories (Lagu & Landry 2011).

The fourth generation of function allocation, which is in its conceptual stage, refers to methods that determine the amount of reallocation needed to affect the workload of the human agent (Lagu & Landry 2011). This is an issue which was not addressed in the first three generations of function allocation.

2.6.3 A function allocation procedure

There are many functions in agricultural vehicles that can be candidates for automation. For research purposes, it is not practical to design experiments with randomized automation of all variables. A model is needed to decide the candidate functions for automation and the level of automation for those functions. In recent studies of adaptive automation, the four-stage model of Parasuraman et al. (2000) has been used as a starting point (de Tjerk et al. 2010). It must be recognized that automation of a function can vary by type (Figure 2-1). Parasuraman et al. (2000) proposed four classes of functions that automation can be applied to: i) information acquisition, ii) information analysis, iii) decision and action selection, and iv) action implementation. Next, it is necessary to identify the level of automation for each class of functions. Evaluations of the system are the next steps. The model proposed two evaluative criteria. The first evaluative criteria consider human performance consequences such as workload, situation awareness, complacency, vigilance decrement, and skill degradation. Secondary evaluative criteria are that the system should be evaluated from the perspective of automation reliability and cost. This framework ends with a determination of the final type and level of automation. If the data from the primary evaluation criteria cover the entire situation that may happen in the system, then it is appropriate to use a dynamically automated system which can provide feedback to the operator and modify the level of automation according to the system situation.

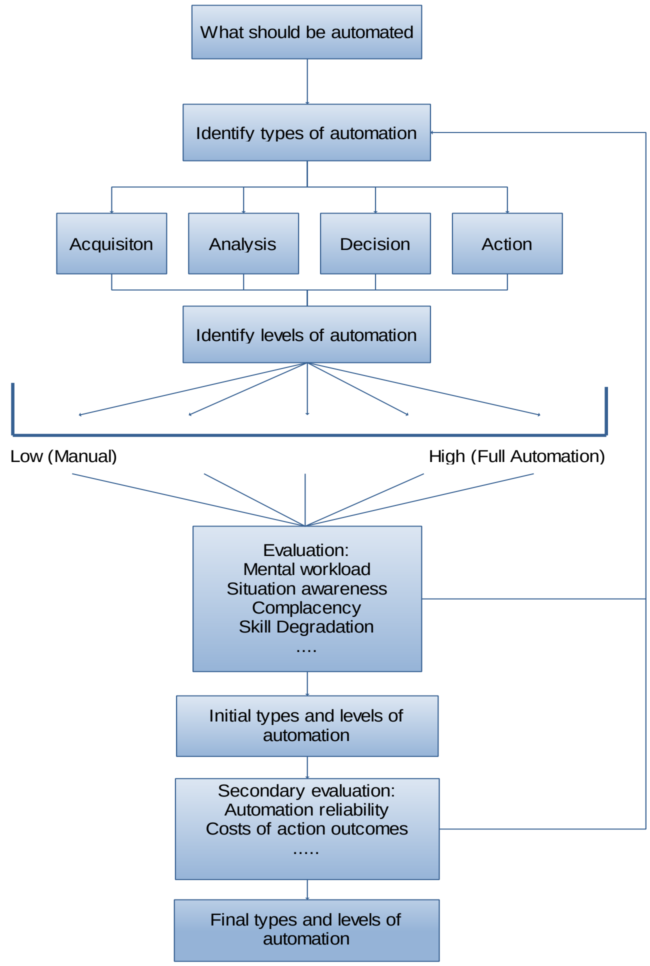


Figure 2-1. The four-stage model of Parasuraman et al. (2000) for function allocation.

2.6.3.1 Degree of automation

In the four-stage model of Parasuraman et al. (2000), four classes of functions have been specified. These classes of functions are derived from the four-stage model of human information processing which consists of sensory processing, perception/working memory, decision making, and response selection. This means that a function is associated with four filters in a hierarchy to be accomplished.

Information acquisition refers to the sensory process of information processing (i.e., detection and registration of data). Automation support in the case of the information acquisition function can be in the form of prioritizing and highlighting some part of the information (Parasuraman, 2000). Information analysis consists of conscious perception and manipulation of processed and retrieved information in working memory (Balfe, 2010) that can be used for predictions. Automation of this function means that the machine can help with prediction of the future state of the vehicle and projection of errors. Decision and action selection indicates the state of choosing from different decision options. At the lowest level of automation, the operator is solely responsible for making a decision. At the highest level, the machine decides and acts automatically. Action implementation refers to the execution of the chosen action. Replacing human hand or voice is the purpose of automating action implementation.

For the automation of functions, a model of level of automation (LOA) must be used. A 10-level automation scale proposed by Sheridan and Verplank (1978, Table 2-1) has been mentioned to best characterize the decision/action selection (Parasuraman et al., 2000). Parasuraman et al. (2007) suggested a more generalized scale of LOA for any type of task in the form of none, low, medium, high, and full automation. Similar scales are proposed by Endsley (1999) and Proud et al. (2003). Utilizing any of these scales may show some constraints. All models of levels for automation start with manual function at the lowest level and end with a highly automated function at the highest level. The number of levels between these extremes can be defined according to the intended complexity of the system and/or the availability of technology. For example, for automation of information acquisition, the lowest level may involve strategies for mechanically moving sensors in order to scan and observe while the moderate levels may have criteria for organization of incoming information such as highlighting some part of the information or having a priority list.

Table 2-1. A 10 level model of automation for decision and action selection (Parasuraman et al. 2000)

High	The computer decides everything, acts autonomously, ignoring the human.
	Inform the human only if it, the computer decides to
	Inform the human only if asked, or
	Execute automatically, then necessarily informs the human, and
	Allows the human a restricted time to veto before automatic execution, or
	Execute that suggestion if the human approves, or
	Suggests one alternative
	Narrows the selection down to a few, or
	The computer offers a complete set of decision/action alternatives, or
Low	The computer offers no assistance: human must take all decisions and actions.

When designing LOA experiments, it may not be possible to assign each LOA to information processing functions in random. For example, if the LOA for information acquisition is none (manual) or very low, accordingly the LOA for information analysis will be very low and assigning a higher LOA may not be possible. Besides, it is mentioned that for different functions there may be unique ranges of automation (Kaber et al. 2005). This means that for different systems, different ranges of automation should be defined.

Due to complications in defining LOA for different systems, researchers have made some recommendations for assigning initial LOA and designing experiments; these include assigning initial LOA for information processing functions and/or merging information processing functions and assuming few levels of automation. In some LOA studies, individual information processing functions were automated in each experimental condition. This condition can be found in studies by Clamann et al. (2002), Kaber et al. (2005), McClernon et al. (2006) and Manzey et al. (2008). Some researchers suggested considering high levels of automation for information acquisition and analysis functions and a moderate level for decision-making, however, a higher level can be used for highly reliable automated systems. It is also mentioned that there is a close affinity between information acquisition and information analysis, and between decision making and action implementation (Parasuraman & Wickens 2008). In the literature, the first affinity is called "information automation" while the latter is called "decision automation" (McGarry et al. 2003).

A study by Galster et al. (2002) on the effects of information automation and decision-aiding cueing on action implementation in a visual search task showed that at higher workloads significant detection yielded in information automation condition. Another study by Dorneich et al. (2001) showed that information automation caused a good balance between the workload of airline dispatchers and their situation awareness while leaving them in complete control of the system.

2.6.3.2 Primary evaluation Criteria

After identification of the function that should be automated and the type and level of automation, it is necessary to evaluate the system for human performance consequences. Human performance areas that can be negatively affected by automation include mental workload, situation awareness, complacency, vigilance decrement, and skill degradation. Although there is some literature about these human performance areas, most of the studies are focused on mental workload and situation awareness. In fact, these factors play a vital role when designing user interfaces and performing system evaluation (Vidulich 2002). In this review of the literature, mental workload and situation awareness are the evaluative criteria that will be emphasized.

2.6.3.2.1 Mental workload

Mental workload is the most important subject of human factor studies. Mental workload reflects perceptual and cognitive demands of tasks on the limited mental resources of humans (Di Stasi et al. 2011). In fact, workload has been used to explain the interaction between the task and the operator who is doing the task (Prinzel et al. 2003a). There have been many research studies on mental workload of operators in human-machine systems (Baldwin et al. 2004; Cantin et al. 2009; de Waard 1996; Desai 1993; Di Stasi et al. 2009; Pauzié 2008a; Veltman & Gaillard 1993; Young & Stanton 1997) and many methods have been developed. Four categories of characteristics can be used to assess mental workload (i.e., subjective measures, physiological measures, primary task measures, and secondary task measures) (Luximon & Goonetilleke 2001). Some authors suggest a three-category classification and put primary and secondary task measures in the category of performance-based or behavioral measures or measures of task performance (Brookhuis et al. 2009; Desai 1993; Veltman & Gaillard 1993). It is suggested that multiple workload measures for a study give greater sensitivity to workload variations than each measure individually (Eggemeier & Wilson, 1991) and the assessment would be more reliable than any measure by itself (Parasuraman et al., 1992). As Cain (2007) noted, different measures may be sensitive to different aspects of workload and not all workload measures are assessing the same thing. A brief description of each category of workload metrics is provided in the following paragraphs.

2.6.3.2.1.1 Measures of task performance

Task performance measures in the driving context are used to assess the ability of drivers to perform tasks accurately on a time-limited basis. The goal is direct assessment of the operator's capability to perform the driving task at an acceptable level (Brookhuis et al. 2009). Reaction time and accuracy of actions are two important components of task performance. Degree of automation, task difficulty, and task type can directly influence these components. Improvement in reaction time was reported when drivers were driving with an autonomous control mode in military semi-autonomous vehicles (Gempton et al. 2013). Another experiment by Sethumadhavan (2009) demonstrated the benefits of high levels of automation in multi-task environments where operators had to perform multiple tasks concurrently. Johnson and Widyanti (2011), in their study on cultural influences on the measurement of subjective mental workload, measured reaction time of subjects in response to a hybrid memory/visual search task. They found that by increasing task difficulty, reaction time was increased.

Measures of task performance can be done directly by measuring primary task performance (i.e., vehicle handling tasks such as steering and car following), or indirectly by means of secondary tasks (i.e., responding to a cell phone). In the case of primary task performance, the numbers of human errors, performance speed, or reaction time can be used as performance measures for laboratory tasks. For real world applications, primary task performance is very task specific. It is stated that primary task performance gives a measure of the overall effectiveness of man-machine interaction. According to de Waard (1996) it is necessary to conduct another measure with primary task performance to reach a valid conclusion. A secondary task refers to any task that is not a normal function of a system (Wierwille & Eggemeier 1993). When the secondary task is added to the primary task, secondary-task measures can be taken.

Kantowitz (1995) used performance measures to assess the workload of heavy truck drivers in a fixed-base truck simulator. In this study, six primary task and four secondary workload measures were assessed. The results showed that most of the primary task measures were not influenced by secondary task, except steering. He concluded that secondary task measures provided effective measures of driver workload. Kantowitz et al. (1996) continued this study with more complex secondary tasks (i.e., using a cell phone to make calls and to read and respond to text messages). The study illustrated that tasks that require reading a message had the greatest impact on driver performance.

Hurwitz and Wheatley (2002) also found the same results as Kantowitz (1995) regarding the negative effect of a secondary task on the steering task. They measured performance of drivers while driving with added auditory or visual monitoring tasks. Their results showed that secondary visual tasks had greater impact on the control of the vehicle than secondary auditory tasks.

2.6.3.2.1.2 Subjective reports

Subjective reports have been used for assessment of the mental workload of operators (Laux & Plott 2007). Self-reports and observer reports are two kinds of subjective reports. Self-reports are given by the operator, as the test subject, to rate his/her experienced mental workload on a task. Here, the questions and rating scale are provided by the researcher. In the case of an observer-report, experts observe and then rate the workload of operators in different test or work situations. Both techniques are simple to do, low cost, and non-intrusive. Self-reports are usually negatively affected by personal interpretations, while the observer reports remove this by featuring strict protocols. Subjective reports are typically used in association with other workload measurement techniques (Sheridan 1991).

An example of self-report workload assessment tool is the NASA Task Load Index (NASA-TLX) rating scale. The NASA-TLX, which was originally designed to assess pilot workload in the aviation domain (Pauzié 2008a), allows operators to rate their experienced workload from six separate dimensions (i.e., mental demand, physical demand, temporal demand, performance, effort, and frustration level) (Laux & Plott 2007). The final outcome from the NASA TLX is an overall workload score based on a weighted average of ratings on the mentioned factors (Yung-Tsan et al. 2009).

The Driving Activity Load Index (DALI) is a revised version of NASA-TLX for the driving task (Pauzié et al. 1995). This tool aims at identifying the origins of the workload of drivers. DALI has been used and validated in various studies related to the driving task, such as secondary task context or diversified driving situations with varying levels of complexity (Pauzié 2008b). It evaluates three workload components of the driving task including perceptual load, mental workload and driver's state. Perceptual load is composed of three factors: visual demand, auditory demand, and tactile demand. The factors comprising mental workload include attention, temporal demand, and interference. Situational stress is an indicator of the driver's state. Considering the mentioned factors, depending on the driving or test condition, up to seven workload factors can be measured. If any factor does not seem to be important to evaluate, it can be omitted from the questionnaire. Petzoldt et al. (2011) used only five subscales of DALI, as the auditory demand and tactile demand were not applicable in their study on the Learning effects in the lane-change task.

2.6.3.2.1.3 Physiological measures

Some physiological measures have shown to be valuable measures of the workload of operators. Several advantages of physiological measures are the ability to achieve continuous measurement with little intrusion, easy measurement of resource capacity, and diagnosis of multiple levels of arousal, attention, and workload. The three most promising physiological measures of mental workload are the electroencephalogram (EEG), event-related potential (ERP), and heart-rate variability (HRV) physiological signal (Kramer et al. 1996; Prinzel et al. 2003b). Experiments by Prinzel et al. (2003b) confirmed the usefulness of these three measures for adaptive automation design. Schrauf et al. (2011) examined the impact of auditory secondary tasks on driving performance through measuring reaction time (as a performance measure) and EEG and HRV (as physiological measures). Both performance measures and physiological measures gave similar results, indicating the usefulness of EEG and HRV for real road driving experiments.

HRV is probably the most used physiological measure in mental workload measurement experiments (Meshkati 1988). Heart rate (HR) and HRV have been shown to be sensitive to changes in physical and mental workloads (Brookhuis & de Waard 2010). According to Mulder et al. (2004), mental efforts are associated with increased HR and decreased HRV. The term ‘heart rate’ describes the frequency of cardiac cycle in a person and is usually shown as the number of heart beats per minute (bpm). Heart rate variability (HRV) refers to variations in heart rate or more specifically variations of intervals between consecutive R peaks of the heart beat waves. Heart beat is under the control of the autonomic nervous system that consists of two components: i) sympathetic nervous system (SNS) and ii) parasympathetic nervous system (PNS). The interaction between SNS and PNS results in variations in heart rate (Rajendra et al. 2006). Under the condition of acute time pressure and emotional strain, the PNS is suppressed and the SNS gets activated. Increased sympathetic activity results in increased heart rate. Conversely, PNS activation slows the heart rate.

To examine the fluctuations in autonomic nervous activity, power spectral analysis of heart rate variability (HRV) has been used instead of traditional cardiovascular measurements (Sato & Miyake 2004). Mulder distinguished three different frequency bands for HRV including a low-frequency area (0.02 - 0.06 Hz), a mid-frequency band (0.07 - 0.14 Hz), and a high-frequency band (0.15 - 0.40 Hz) (Mulder et al. 2004). Some authors stated these ranges as the very low frequency band (VLF, 0.02 - 0.06 Hz), the low frequency band (LF, 0.07 - 0.14 Hz, known as 0.1 Hz component of HRV), and the high frequency band (HF, 0.15 - 0.40 Hz). The aforementioned frequency ranges also have been slightly altered by some researchers. In spectral analysis, LF variations reflect sympathetic and parasympathetic activities, while HF variations reflect parasympathetic activity. LF/HF ratio, therefore, indicates overall balance between SNS and PNS. Indices of SNS (I_{SNS}) and PNS (I_{PNS}), as normalized powers of LF and HF have been used as measures of autonomic nervous activity in response to mental and physical demands (Garde et al. 2002). These normalized powers of LF and HF are calculated using total power (TP, 0 – 0.4 Hz) and VLF as follow: I_{SNS}=LF/(TP–VLF) and I_{PNS}=HF/(TP–VLF).

It has been reported that the 0.1 Hz component of HRV follows a decreasing trend as mental load increases (Ramon et al., 2008), so it has been used in different studies to assess mental workload demands. A study by Dey and Mann (2010) using HRV measurement showed that the mental workload (0.1 Hz component of HRV) of an agricultural sprayer operator was increased when driving with a lightbar compared to an auto-steer navigation device. In another study, Nickel and Nachreiner (2003) found that the 0.1 Hz component of HRV was only able to differentiate mental workload between activity and rest periods where the differences between mental workloads were relatively large.

Although the HRV technique has shown correlations with mental effort in many studies, contradictory results have been reported (Li et al. 2013). Nickel and Nachreiner (2003) argued against using HRV as a measure of mental and cognitive workloads. Results from their study did not support acceptable sensitivity and diagnosticity of the 0.1 Hz component of HRV as an indicator of mental strain. Engström et al. (2005) tried to differentiate the effects of visual and cognitive load on driving performance and driver state in motorway driving. They examined an in-vehicle information system with different levels of difficulty. No main effects were found on HRV in a moving base simulator.

2.6.3.2.2 Situation awareness

In the past two decades, there have been many studies on situation awareness (Durso et al. 2006; Durso & Sethumadhavan 2008; Endsley 1995, 1996, 1999; Endsley & Garland 2000; Kaber & Endsley 2004; Salmon et al. 2009; Sonnenwald et al. 2004), and the term situation awareness has become a vernacular to human factors research (Wickens 2008). Endsley (1988) describes situation awareness as “the detection of the elements in the environment within a volume of space and time, the comprehension of their meaning, and the projection of their status in the near future” (Endsley & Robertson, 2000). In fact, situation awareness is an operator's dynamic understanding of 'what is going on' (Salmon et al. 2009).

In human-machine systems, the situation dynamically changes with time, so there will be some times when the operator does not get sufficient feedback. This happens when automation takes over the situation, and as a result, the operator's situation understanding deteriorates (Bye et al. 1999). A hypothetical change of situation understanding is shown in Figure 2-2. Low situation understanding means lower system performance. This is a major concern when designing human-machine systems. There have been a variety of efforts to find solutions to keep human situation awareness at an acceptable level. Providing real-time feedback (Walker et al. 2006) and metacognitive-strategy training (Soliman & Mathna 2009) have been shown to improve situation awareness. It should be noted that metacognitive strategy training involves strategies for learning about the learning process. It includes "explicit control of learning, planning and selecting strategies, monitoring the progress of learning, correcting errors, analyzing the effectiveness of learning strategies, and changing learning behaviors and strategies when necessary" (Soliman & Mathna 2009).

In the driving domain, Stanton, Dunoyer, and Leatherland (2011) defined situation awareness as "understanding the relationship between the driver's goal, the vehicle states, the road environment and infrastructure, and the behavior of other road users at any moment in time." In fact, the notion of situation awareness here describes the ability of the driver to combine longer-term goals, such as driving toward a destination, with shorter-term goals, such as avoiding collisions, in a real-time process (Sukthankar, 1997). A similar definition of situation awareness can be imagined for driving agricultural vehicles. However, the task differences between driving a car and driving an agricultural vehicle on short-term and long-term goals should not be neglected. Therefore, the term situation awareness in the domain of driving agricultural vehicles must include understanding the relationship between the operator's goals, the state of the agricultural vehicle and implement, and the field environment; and understanding the behavior of other operators when there is more than one vehicle working in a field.

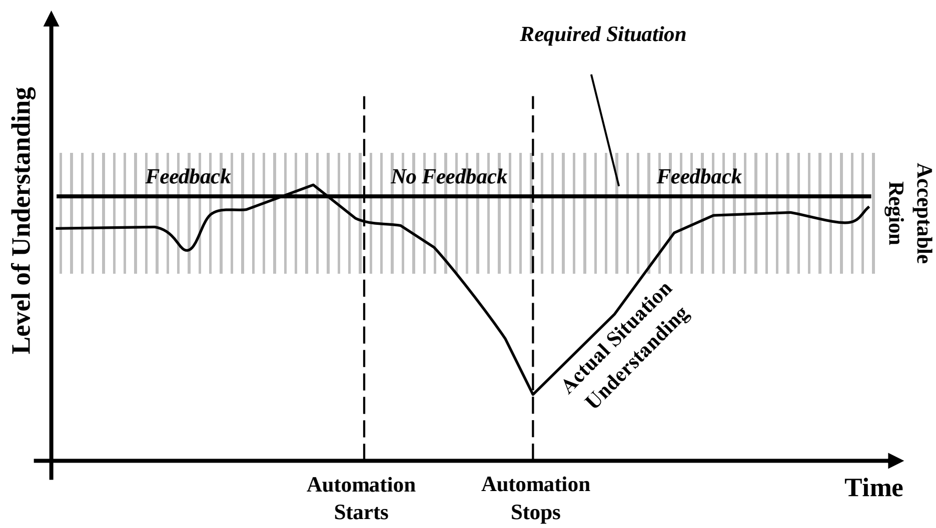


Figure 2-2. Situation understanding with and without feedback (Bye et al. 1999)

There are numerous methods for measurement of situation awareness (Endsley et al. 1998; Strybel et al. 2007), which we can categorize as follows: freeze probe recall techniques, real-time probe techniques, subjective reports (post-trial subjective rating techniques, observer rating techniques), performance measures, and process indices (Salmon et al. 2009). A description of these methods is provided in the following paragraphs.

2.6.3.2.2.1 Freeze probe techniques

This technique directly assesses the situation awareness of operators in a simulated task environment during task performance (Salmon et al. 2009). The researcher randomly freezes a task, turns all screens and displays off, and provides the participants with a set of queries regarding the situation at the time it was frozen. The answers to queries, which are based upon their knowledge and understanding of the situation at the point of the freeze, are compared to the state of the system. At the end of a trial, an overall situation awareness score is calculated. Although this technique offers direct, subjective measurement of situation awareness and removes the issues of data collection after tests, it has been criticized for being difficult to use for real-world activities and of being intrusive of task performance (Salmon et al. 2009).

The situation awareness global assessment technique (SAGAT) developed by Endsley (2000) is the most popular freeze probe technique (Salmon et al. 2009). SAGAT offers questions based on three levels of situation awareness for comprehensive assessment of the situation awareness requirements of operators (Endsley & Garland 2000). Experiments by Kaber et al. (2006) using SAGAT to assess the effectiveness of adaptive automation of air traffic control tasks showed that when automation was applied to information acquisition and action implementation, the performance was significantly higher compared to automation of cognitive functions, specifically information analysis. Experiments by Walker et al. (2006) suggested the feasibility of SAGAT for driving simulator studies. They examined the situation awareness of drivers through SAGAT in a simulator by providing non-visual vehicle feedback. Their study confirmed the importance of providing feedback to the drivers for coupling the driver to the dynamic driving environment. Another simulator study by Soliman and Mathna (2009) showed that metacognitive training strategies can enhance the situation awareness of drivers.

2.6.3.2.2.2 Real-time probe techniques

This technique for measurement of situation awareness is used when the operation cannot be interrupted. This technique provides questions asked directly to the operator during ongoing operations rather than during periods when the system has been frozen (Jones & Endsley 2000). Situation Present Assessment Method (SPAM), developed by Durso et al. (1996), is a real-time probe technique that removes the problem of relying on memory (i.e., off-line query methods) (Durso et al. 2006). In an air traffic control study, Durso et al. (1998) could not find evidence of the validity of real-time probes for measuring situation awareness. Studies by Jones and Endsley (2000; 2004) suggest more research for evaluating the utility of real-time probes for assessing situation awareness. Zhang et al. (2009) used a real-time probe technique to study the effect of the age of the driver on situation awareness in hazardous situations using a driving simulator. Results showed that older drivers had lower overall situation awareness than younger drivers.

2.6.3.2.2.3 Subjective reports

These methods, similar to the subjective measures of workload, require that either operators provide self-ratings of situation awareness, or observers rate an operator's situation awareness (Strybel et al. 2007). Self-rating techniques are non-intrusive, easy, quick and low cost (Jones 2000). Observer-rating techniques offer similar advantages as self-rating techniques and they are commonly used 'in-the-field' due to their non-intrusive nature. Self-rating techniques are criticized for a number of reasons, including the problems associated with post-trial data collection. There are also some concerns about the validity of observer rating techniques (Salmon et al. 2009).

The Situation Awareness Rating Technique (SART) developed by Taylor (1989) is a widely used subjective report technique for the measurement of situation awareness (Salmon, Stanton, Walker, Green, 2006). SART focuses on generic, overall task characteristics rather than the specific elements related to the task (Salmon *et al*. 2009). It can measure up to 10 dimensions of SA. These dimensions can be categorized in three distinctive groups: i) *demand on attentional resources*, ii) *supply of attentional resources* and iii) *understanding*. *Demand on attentional resources* includes questions regarding instability, complexity, and variability of the situation. *Supply of attentional resources* reflects arousal, concentration, division and spare mental resources when dealing with the situation. Finally, *understanding* of a situation depends on the quality and quantity of the information provided as well as familiarity with the situation. Values for each of the three categories can be derived by getting the averages of responses to corresponding queries. Finally, the combined rate for situation awareness can be inferred by the following formula (Jones, 2000):

SA = U – (D – S)

Where:

SA = Situation awareness

U = Understanding of the situation

D = Demand on attentional resources

S = Supply of attentional resources

2.6.3.2.2.4 Performance measures

Performance measures of situation awareness are indirect measures of situation awareness. They refer to the measurement of relevant aspects of the performance of operators while performing the task. The aspects that must be measured are task dependent (Salmon et al. 2009). For example, Gugerty (1997) measured hazard detection, blocking car detection, and crash avoidance during a simulated driving task as performance measures of situation awareness when assessing the car-driving task.

2.6.3.2.3 Mental workload and situation awareness relationship

It has been stated that mental workload and situation awareness are clearly distinct, but are also intricately related to one another (Vidulich & Tsang 2012). Various researchers have provided conceptual frameworks illustrating the dynamic interaction of mental workload and situation awareness (Durso & Alexander 2009; Vidulich & Tsang 2012; Wickens 1996). Wickens et al. (2010) stated that with a higher degree of automation, hypothetically mental workload is reduced, but situation awareness also declines. Meta analyses by Onnasch et al. (2014) and Wickens et al. (2010) suggest tradeoffs between levels of situation awareness, mental workload, and degree of automation for optimal task allocation and task performance. Vidulich and Tsang (2014) argued that mental workload and situation awareness could support each other or compete with each other, so knowing about only one of them is often not sufficient for assessing task allocations, interface design, or similar features. Therefore, for allocation of functions in semi-automated systems, individual measurement of workload and situation awareness is critical.

2.6.3.3 Secondary evaluative Criteria

Automation reliability and costs of decision/action outcomes are considered as secondary evaluative criteria in the four-stage model of Parasuraman et al. (2000). The automation must be reliable in order to maintain its benefits on mental workload and situation awareness. Masalonis (2003) stated that, in different situations, the reliability of decision support automation can vary predictably. The results could be inappropriate trust in automation and performance decreases. For estimating the reliability of automation of a system, several procedures have been proposed including fault and event tree analysis (Swain 1990) and various methods for software reliability analysis (Parnas et al. 1990).

Assessing the proper level of automation for decision automation requires additional consideration of the costs associated with incorrect or inappropriate decision and action. These costs or consequences vary by the type of the action that the human or automated system takes. We can evaluate the risk associated with a decision outcome by multiplying the cost of an error by the probability of the occurrence of the error.

2.7 Simulator: A tool for human factors studies

2.7.1 Introduction

Collecting data for task analysis can be done in actual situations or in a simulated situation in the laboratory (Kieras 1997). Due to reasons such as difficulty and cost of experiments in actual situations, and to have control over the variables under study, researchers of man-machine systems have relied on the use of simulators (Fitts 1951). Driving simulators have been developed and used in the automobile industry since at least the mid-1960s (Weir 2010). A variety of training, research, and design applications have been supported by driving simulation. As a research tool, they provide unique opportunities in terms of experimental control, flexibility, cost, and safety. Driving simulators have been widely used to study various aspects of driving automobiles including human factors aspects of driving (Cantin et al. 2009; Koutsopoulos et al. 1995; Rakauskas et al. 2004), development of in-vehicle human–machine interfaces (Weir 2010), human perception and control (Brookhuis et al. 2009; Kemeny & Panerai 2003), clinical research (George 2003), and the design of vehicles and roadways (Kawamura et al. 2004).

The development and usage of driving simulators of agricultural vehicles date back to the early 1970s. Zander (1972) described a man-task system simulator for studying the effects of factors such as vibration and temperature on the operators of combine harvesters (Preston 1979). Sjøflot (1976) developed a driving simulator for studying the operation of a tractor-forage harvester system (Kaminaka & Fortis 1983). The Federal German Berlin University Group also had a tractor driving simulator which permitted the whole tractor rather than just the seat to be used for studies inside the laboratory (Preston 1979). Kaminaka and Fortis (1983) constructed a cost-effective digitally controlled tractor simulator to study the visual monitoring performance of tractor operators. Wilkerson et al. (1993) described a tractor simulator to simulate roll-overs. All of the mentioned simulators are obsolete considering the technology that is available today; recent technological advancements in computer hardware and software and changes that have occurred in tractors demand new tractor driving simulators that can mimic the current real situation of such vehicles. In an attempt to study ergonomic issues of tractor driving, Karimi et al. (2008) constructed a tractor driving simulator using a cab salvaged from an old tractor. They used the tractor simulator to study some aspects of driving a tractor. They found that the motion cues have significant effect on driving an agricultural vehicle in parallel swathing mode. They also studied the role of torque feedback on the steering wheel of an agricultural vehicle during common field operations of straight line driving and monitoring of a rear-mounted machine (Karimi & Mann 2009).

2.7.2 Driving simulation, advantages and shortcomings

A driving simulator makes it possible to operate vehicles with no real movement, but in realistic conditions (Käppler 2008). In fact, a driving simulator is a system that provides an intelligent environment in which a human driver can perceive and control the operation of a virtual vehicle. If the driving simulator is to reflect real situations, it must cause the same driving behaviours from drivers as they exhibit in real-world driving. To achieve this goal, the driving simulator must have the same appearance and dynamics as the real vehicle and provide the same information to the driver. It also must provide the same tools to input the necessary control demands. Some driving simulators provide only visual feedback, but most high fidelity driving simulators provide motion, haptic, and auditory feedback which allow the driver to interact with the vehicle and the environment in a multisensory fashion (Kemeny & Panerai 2003). Extensive research has shown that, depending on the driving task being simulated, non-visual cues are necessary to provide realistic simulation (Siegler et al. 2001; Steele & Gillespie 2001).

For research purposes, experimental control is the greatest advantage offered by driving simulators. Using driving simulators, it is possible to control many extraneous variables that cannot be controlled in real driving. It is also possible to separately control independent variables as desired and to run several experiments under equal experimental conditions. Much less planning is needed for conducting experiments in a driving simulator. It is also much less costly to conduct experiments with a driving simulator compared to experiments in an instrumented car in a real environment. The safety of the driver in the test is another important advantage of driving simulators. This factor is of most significance when studying issues such as driver fatigue or driving during low visibility conditions. It is also much easier to measure driving performance variables and other parameters, such as physiological and psychological responses of the driver, in a driving simulator than in a real vehicle (Horiguchi & Suetomi 1995).

Despite the advantages, driving simulators have certain shortcomings. No driving simulator can perfectly reproduce the real driving experience. Models of vehicle dynamics and environmental disturbances can be made increasingly accurate, but can never be perfect. Providing visual feedback that has the same field of view, resolution, and depth cues as those of a real visual scene is extremely difficult, if not impossible. In addition, even in the most advanced driving simulators, certain motion cues are not possible to render, because no driving simulator has an unlimited motion range. Direct rendering of simple vehicle maneuvers (such as a long brake) requires large motion systems that are unrealistic. Engineers have developed special techniques such as motion washout filtering, tilt coordination, and motion scaling that can render most vehicle motions, but these techniques do not completely resolve the existing problems. Transport delay is another major issue; there is always a delay between the subject’s action and the simulator’s response. This is due to the time required for the acquisition of the subject’s commands, computation of the appropriate response, and the delay in the visual and motion subsystems. Not only should these delays be small, but also all simulator subsystems should be synchronized, a requirement that is difficult to achieve (Horiguchi & Suetomi 1995; Kemeny & Panerai 2003).

Chapter 3

STUDY OBJECTIVES AND FRAMEWORK

3.1 Objectives

In the literature review, several problems associated with human-machine systems have been discussed and methods for investigating those problems and lessening their effects have been explained. It was mentioned that there are only a few studies on agricultural human-machine systems in which some critical events have been considered (i.e., roll-over accident or effects of ambient light condition on driving agricultural vehicles). Driving a semi-autonomous agricultural vehicle on a field has been defined as a monotonous task that requires the operator to repeat some monitoring and physical tasks for several hours. This raises concerns about the human factors issues associated with operating agricultural vehicles. System safety and efficiency will be ensured if the operator-vehicle interaction is designed based on human factors criteria.

This study was designed to study the effects of automation on some human factors in agricultural vehicles. Considering that there are diffrent types of agricultural vehicles that could be studied, a tractor air-seeder system was selected as a case study. The reason for selecting this system is the availablity of some information on its driving tasks from a previous study. The objective of this study was twofold: i) to identify tasks associated with the operation (functional analysis) of an agricultural semi-autonomous tractor with a mounted air-seeder system, and ii) to investigate the effects of task automation of air seeder control and monitoring on two human performance parameters: mental workload and situation awareness.

The lack of basic information for this automation study, as was stated, requires running an in-depth study to achieve the goals. This means there are many variables that can be evaluated in experiments. Applying these variables in an agricultural vehicle worth approximately half a million dollars is a significant challenge. A simulator study eliminates these problems and provides control over the variables under study. Thus, research will be completed with the use of the tractor driving simulator (TDS) located in the Agricultural Ergonomics Lab, Department of Biosystems Engineering, University of Manitoba. This simulator is the only tractor driving simulator in Canada for studying human factors issues in agricultural vehicles.

3.2 Study framework

In this study I followed the straight-forward procedure for function allocation suggested by Parasuraman et al. (2000). As discussed in the literature review, the focus of their model is to apply different degrees of automation to information processing subtasks of a particular task. In order to adapt the model in this study, the first step included identification of tasks associated with the operation of agricultural vehicles. This required observing real-world operation of such vehicles as well as interviewing operators. After this stage, it is possible to discuss the tasks and analyze them. The output from this stage is a graphical representation of a hierarchy of tasks (see Figure 4-1 and Table 4-1).

After task analysis, independent variables of the study were selected and the procedure for assigning levels of automation for each variable is discussed. As the first evaluation criteria, mental workload and situation awareness of operators is assessed at different task automation levels. Mental workload and situation awareness measurement techniques used in this study are discussed in Chapter 5. Due to some technical limitations associated with secondary evaluative criteria such as time constraints of automation reliability and trust, secondary evaluation will not be performed in this study. A preliminary experiment was completed before proceeding to main experiments in order to evaluate the experimental techniques as well as durations of the experimental and training blocks.

Chapter 4

TASK ANALYSIS

4.1 Introduction

As was mentioned in Chapter 3, a tractor-air seeder system (TAS) was chosen as a case study. Before proceeding to the experimental stage of the research, it was necessary to identify all of the tasks that are associated with the operation of such a system. Therefore, a task analysis method was selected based on recommendations in the literature.

4.2 TAS task analysis assumptions

TAS task analysis was completed for a scenario of seeding in a straight line in the field. The reason for selecting this scenario was that, for the economy of operation, driving agricultural vehicles on fields usually occurs in parallel straight paths. As Han et al. (2013) states "an optimal coverage path algorithm can enable a vehicle to effectively travel across a field by following a sequence of parallel paths with fixed spacing". This strategy aims to minimize gaps (i.e., uncovered surfaces between paths) and overlaps (i.e., surfaces already covered).

According to the work scenario, the task analysis focused on the on-field operation tasks. A conventional semi-autonomous TAS was considered for analysis because this represents the equipment that is currently being used by farmers. In this system, the introduction of electronics has led to a proliferation of controls in the tractor cab, which significantly modifies the way users must manage their working task. The use of electronics to manage the attached air seeder allows the operator to perform calibrations, settings, and necessary adjustments while inside the tractor.

4.3 Task analysis procedure

The information used for the task analysis was gathered from the literature and site visits. The information about driving functions and relevant terminologies was primarily obtained from previous studies from the car-driving domain. On the other hand, the information on tasks related to air seeder operation was obtained by reviewing the behavior of TAS operators in real situations. A total of 2 operators were observed in Manitoba (in ride-alongs by the author) and 13 operators were observed in Saskatchewan and Alberta (reported in Karimi et al. 2012). Ride-alongs facilitated the author's familiarity with the TAS systems being used in the Canadian prairies. As a passenger, the author had the chance to observe the operation and video tape it for further review. The author also operated a TAS under supervision of the operator of the vehicle. Statistics related to TAS tasks in this chapter are from the study performed by Karimi et al. (2012). After collecting information, the operating tasks of the TAS were discussed by the research team in the Agricultural Ergonomics Laboratory (i.e., the author, the author's advisor, and a post-doctoral fellow) to select the functions and tasks to be included in the task analysis.

After selecting the appropriate functions and tasks, it was necessary to represent the activities within tasks. A technical approach was used, as function allocation and interface design seemed to be the major human factors issues related to semi-autonomous agricultural vehicles. As was stated in the literature review, hierarchical task analysis is a proper way to represent task activities and to record task knowledge. This method is a useful tool to learn about the structure of different tasks. Hierarchical task analysis involves identifying the overall function in a top-down fashion to include the task, subtasks and the conditions under which they should be completed to achieve that function (Annett 2003).

The analysis consisted of decomposing the tasks into elementary units and organizing them into three levels. The hierarchical task listings in this task analysis represented the main tasks in the first level, subtasks in the second level and major task activities in subsequent levels. After developing a hierarchical task list, the research team discussed it again with the subsequent results being reflected in the final task list. Once main tasks were identified, the subtasks necessary to carry out each task were developed. Subtasks can be described at varying levels of detail. In order to keep the analysis within reasonable bounds, the following rules were used to decide when the tasks had been broken down to an adequate level of detail: i) tasks would be described by as few subtasks as could reasonably achieve the intended purpose of the task, and ii) tasks would be divided into subtasks and task activities only to the level that the model for type and level of automation could be used. By using these rules, tasks, subtasks, and task activities were identified and arranged in a hierarchical fashion.

4.3.1 A description of tasks

Tractor driving tasks are analogous to driving tasks of on-road vehicles. Driving functions for on-road vehicles are divided into two categories: pre-drive and drive (Wheeler et al. 1996). Pre-drive tasks are the tasks that the driver performs before starting the vehicle. A set of subtasks that should be done prior to driving are: planning the operation, inspection of the vehicle, and start-up. Driving on-road refers to the driving subtasks, such as monitoring engine operation and vehicle condition or changing lanes, when the operator is driving the vehicle on a public road.

The pre-drive task of a TAS includes the same subtasks as for on-road vehicles, but with different considerations. Planning the operation in a TAS is usually more challenging. It comprised of a planning strategy to get from the station to the field, as well as planning the economical field coverage path. Moreover, the attached implement includes many components that need to be inspected along with the tractor. The (on-road) driving task of the TAS also has slight differencs with most on-road vehicles. In this system, the vehicle travelling speed is slower than the normal speed of traffic. It is usually necessary to display a slow-moving vehicle emblen to indicate the slow speed of the vehicle in order to warn other road users. Furthermore, the vehicle is considered over-sized which means that some cautions should be exercised while driving on-road.

In case of operation of the TAS on a field, there were some differences of opinion among the members of the research team with respect to categorizing tasks associated with operating an air seeder system as on-field operation may show differences with regular driving tasks. It was disputed whether the *TAS operation in the field* task should be counted as an ancillary task for the driving task or whether it should be categorized as a separate task. Considering the differences between on-road driving and on-field operation of agricultural vehicles, discussed in section 2.5, the research team ultimately decided that it should be a different category. For a TAS system, the seeding task was identified as the highest-level task. Hence, the pre-drive, drive on-road, and operation in the field were determined as the first level tasks or subtasks for seeding (Figure 4-1). In this hierarchy, operation in the field, which is the main subject of this study, addresses the subtasks of operation of the TAS in the field. The subtasks of this task include those associated with driving and operating the air seeder. In this case, the operator devotes a great portion of his/her attention to monitoring and control of the air seeder.

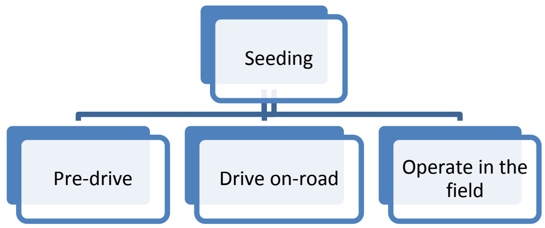


Figure 4-1. Goal and main task level of the TAS hierarchical task analysis.

After developing a hierarchy of high-level tasks, decomposition into subtasks was completed to determine the task activities related to each subtask. Terminology used in the literature was used for naming of tasks where possible. A list of subtasks for operation in the field is provided in Table 4-1. The hierarchical task desription of TAS operation and monitoring task and the task goals are shown in Figure 4-2.

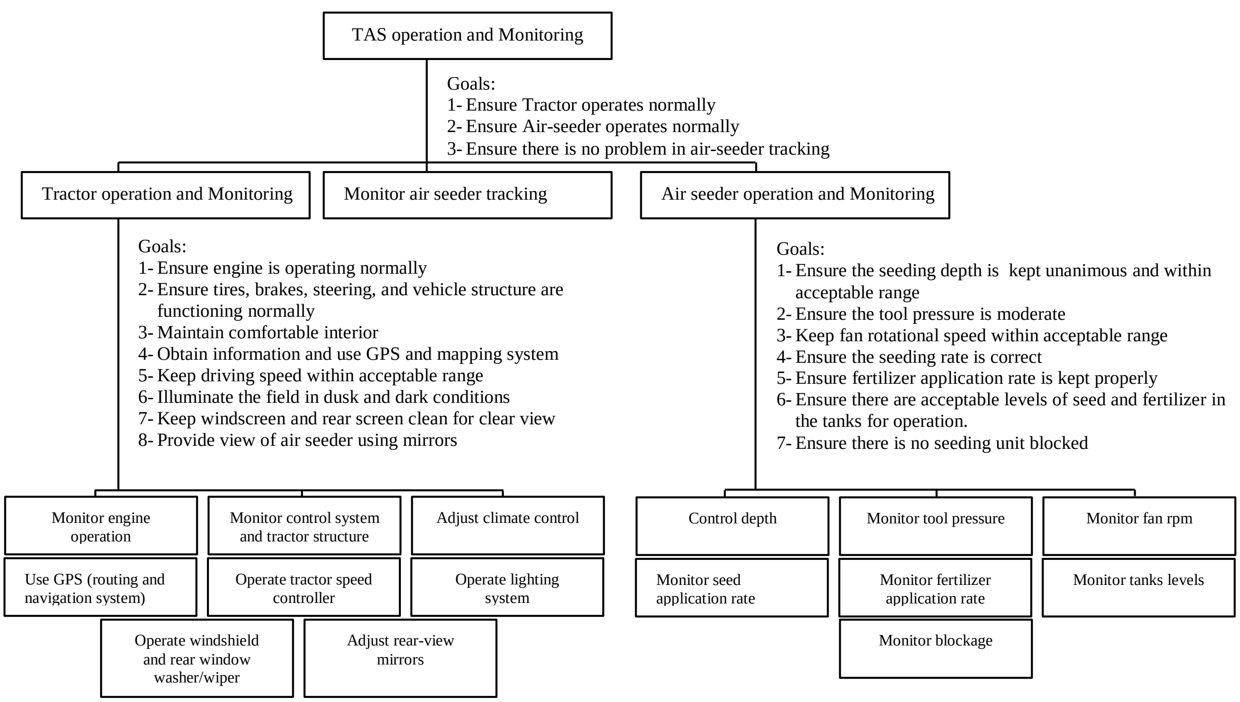


Figure 4-2. A hierarchy of TAS operation and monitoring tasks

Table 4-1. The list of subtasks for operation in the field

Operation in the field task	Subtasks	Task activity
Navigation and routing	Way-finding	Identify Present Location
		Follow planned route
	Route modification	Identify need to correct or change route
		select new route or decide the correction direction
		Execute the modification
TAS operation and monitoring	Tractor operation and monitoring	Monitor control system and tractor structure
		Adjust climate control
		Operate tractor speed controller
		Monitor engine operation
		Use GPS system (Routing and Navigation system)
		Operate lighting system
		Operate windshield and rear window washer/Wipers
		Adjust rear-view mirrors
	Monitor Air seeder tracking	
	Air seeder control and monitoring	Depth control
		Monitor tool pressure
		Monitor fan rpm
		Monitor seed application rate
		Monitor fertilizer application rate
		Monitor tank levels
		Monitor blockage
Control	Speed control	Identify difference between current and desired speed
		Adjust throttle or brake to control speed
		Verify adjustment of speed
	Position control	Identify difference between current and desired lane
		Adjust steering wheel to compensate
		Verify adjustment of lane position
Guidance and manoeuvres	Manoeuvring	Identify present speed and position
		Identify distance to turn point
		Adjust speed and position
		Signal turning manoeuvre near turn point
		Execute turning manoeuvre
	Hazard observation	Estimate hazard potential in the field
		Monitor headway route and surroundings
		Estimate hazard potential to vehicle
		Execute speed and position control
		Execute driving manoeuvre to compensate for hazard
Reacting to emergencies	Detect emergency condition	
	Diagnose situation	
	Determine action required	
	Take appropriate action	

4.3.2 Applying task analysis results to the simulator

The outcome of the task analysis was used to incorporate necessary features in a tractor-driving simulator (TDS). A summary of the outcome of the task analysis for operating a TAS in the field is provided in the following sentences. Operators use a GPS-guidance system as the main source of information for steering the TAS. Operators allocate anywhere from 10-50% of their time to controlling the air seeder (Karimi et al. 2012). They also scan other displays in the TAS cab, including: i) a GPS and mapping system, which shows a bird's-eye view of the field with the tractor's position in the field and the place seeding is being performed, and ii) an application display, which provides such information as the field area covered, forward speed, seeding depth, seed and fertilizer application rates, fan rotational speed, and the amount of seed and fertilizer in the air seeder tank. According to Karimi et al. (2012), with an auto-steer system, drivers spend 30%, on average, of their time looking at the air seeder display.

During normal operation of the tractor-air seeder system, all of the air seeder parameters are expected to stay within acceptable ranges. Due to working conditions, these parameters sometimes move away from the acceptable range, thereby causing system errors. Depending on the level of automation, either the operator or the automated system could be responsible for diagnosis and management of the errors. In the case of TAS steering, if an-auto steer system is engaged, the responsibility of avoiding gaps and overlaps between paths will be given to the automated system. Otherwise, the operator is expected to use the GPS guidance system to perform the task.

Based on the information obtained from task analysis, an application display was installed in the simulator. The display was divided into two sections. One section was dedicated to a mapping system that showed the area covered and position of the TAS in the field. The second section was used as the air seeder information display. Configuration of the air seeder information display was selected based on a study conducted by Karimi et al. (2011). Figure 4-3 shows the air seeder display and the mapping system. Item 1 in Figure 4-3 represents the forward travelling of the tractor (green stripe) in the field and item 2 shows the forward speed of the tractor together with its acceptable range. Items 3 to 11 are air seeder parameters. Item 3 shows the tank level (i.e., percentage of seed or fertilizer remaining in the air seeder tanks). Items 4 and 5 show the seed and fertilizer application rate together with their acceptable ranges. Items 6 and 7 display fan rotational speed and tool pressure, respectively, together with their

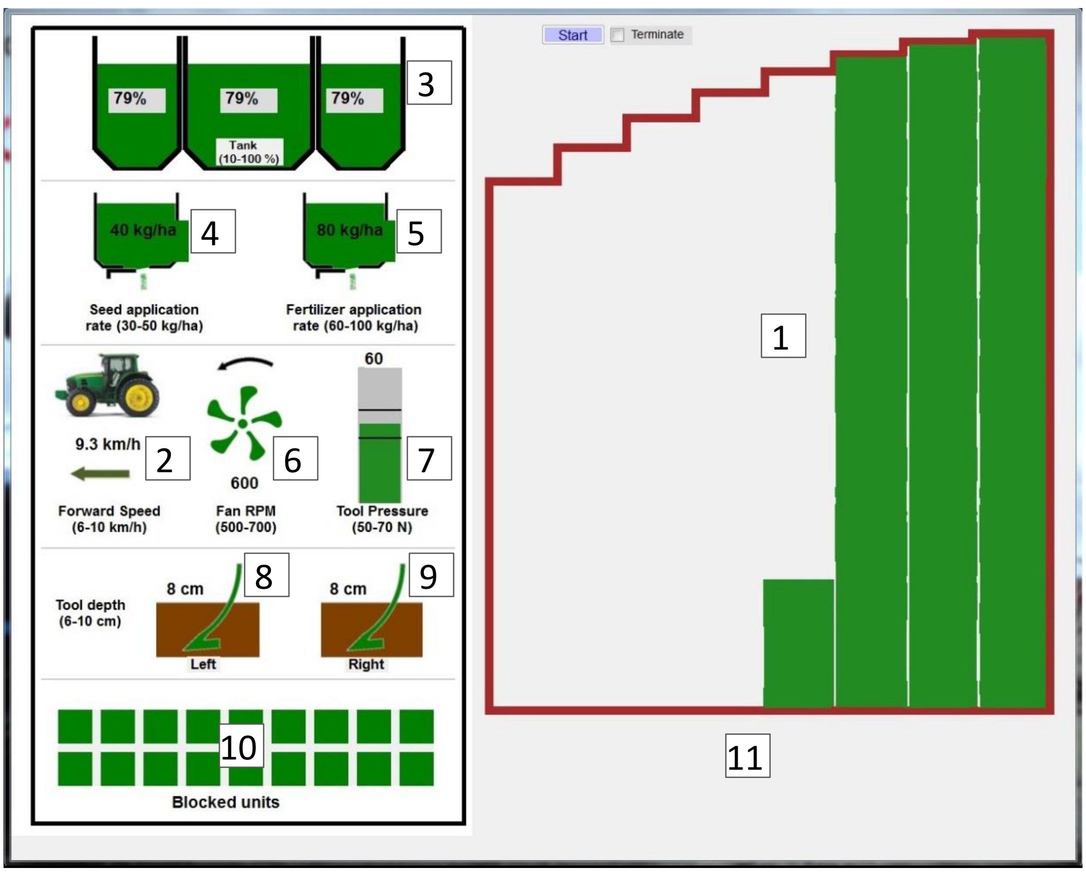


Figure 4-3. The air seeder display and the mapping system.

acceptable ranges. Items 8 and 9 display the working depths of tools from the left and right edges of the air seeder together with their acceptable ranges. Item 10 shows whether any seed distribution tubes are blocked. Finally, item 11 is a message box that is used to provide information regarding air seeder parameters. It should be noted that the message box was not included in the initial configuration of the display. It was added after defining the variables of the study by shrinking the field map on the display.

Chapter 5

EXPERIMENTAL DESIGN

5.1 Introduction

After the task analysis, a simulator experiment was designed and completed in two stages. First, a pilot study with a few participants was completed. The pilot study was used to evaluate the proposed experimental procedures and to inform the final experimental design. After this stage, a full-scale experiment was conducted. The procedure was identical for both experiments as the pilot study confirmed that the methodology was appropriate. The differences between the pilot study and the full-scale experiment will be described in this chapter. The study was in accordance with an ethics protocol that was approved by the University of Manitoba Education/Nursing Research Ethics Board (Protocol Reference Number: E2012:066, Appendix A). The following sections detail the processes and procedures involved in this investigation.

5.2 Participants

The experiment required the use of human participants. Volunteers were recruited from the University of Manitoba graduate and undergraduate student populations via postings on campus. For the preliminary experiment, 10 subjects (9 male, 1 female) were recruited. Age of participants ranged from 20 to 35 yr (M = 26.5, SD = 5.38 yr). They were from for different ethnicities. All participants were required to have at least one year of car-driving experience. In terms of tractor driving experience, two participants were highly experienced (i.e., more than 10 yr) while five participants had no experience. Participants were compensated monetarily at the end of testing for their 2 h of participation.

For the full-scale experiment, *A-priori* analysis using G-Power version 3.1 (Faul et al. 2007) was conducted in order to determine appropriate sample size, and expected power values. The results from the pilot study were not used in this sample size calculation and power analysis since the participants of full scale experiment were selected from a different population. Calculations were based on medium effect size defined by Cohen (1988), (i.e., f = 0.25). The necessary total sample size calculated by G-Power was 24 subjects given the following parameters of a 2 (between-subjects) by 5 (within-subjects) experimental design: f = .25, α = 0.05, and *power* = .85.

30 university students (28 male, 2 female) with at least one year (season) of agricultural tractor driving experience were recruited. They were mostly from farm families. Mean tractor driving experience was 7.7 yr (minimum 1 yr, maximum 14 yr) meaning that all participants had been exposed to the driving task for many hours and were considered to be experienced tractor drivers. Only four of them had less than 5 yr of tractor driving experience. Participants were young, ranging in age from 18 to 25 yr (M = 20.93, SD = 2.05 yr). In terms of ethnicity, there were 28 Caucasian and 2 Chinese participants. None of them had prior experience with the current version of the TDS. Only one participant had participated in a study with the previous version of the TDS. Participants were compensated monetarily upon completion of the experiment.

5.3 Apparatus

An updated version of the TDS was used in this study. A complete description of the previous version of the TDS (Figure 5-1) can be found in Karimi (2008) and Dey and Mann (2011). Figure 5-1 illustrates the tractor cab and interface of the previous TDS.

Figure 5-1. The first version of TDS.

Compared to the previous version, a contemporary tractor cab was implemented in the new simulator. The interface of the new cab was quite different than the previous version in terms of design, driver comfort, and control layout. In terms of the visual scene projection, the three flat screens were replaced by a curved projection screen to provide a naturalistic scene view. Tractor operating noise was another addition to the previous version, as it could have an effect on driving behavior. Actual tractor noise was recorded, at different speeds of operation, during ride-alongs with air seeder operators during the spring of 2012. The recorded noise was incorporated to the simulator code in such a way that it could vary by tractor speed. A pair of speakers was located inside the cab to play the noise to drivers. The layout of the simulator is shown in Figure 5-2. Components of the simulator can be seen in Figure 5-3.

The simulator was programmed to introduce system errors at random times because machine performance is not expected to be optimal at all times. Test participants were expected to monitor air seeder parameters (provided on the air seeder information display

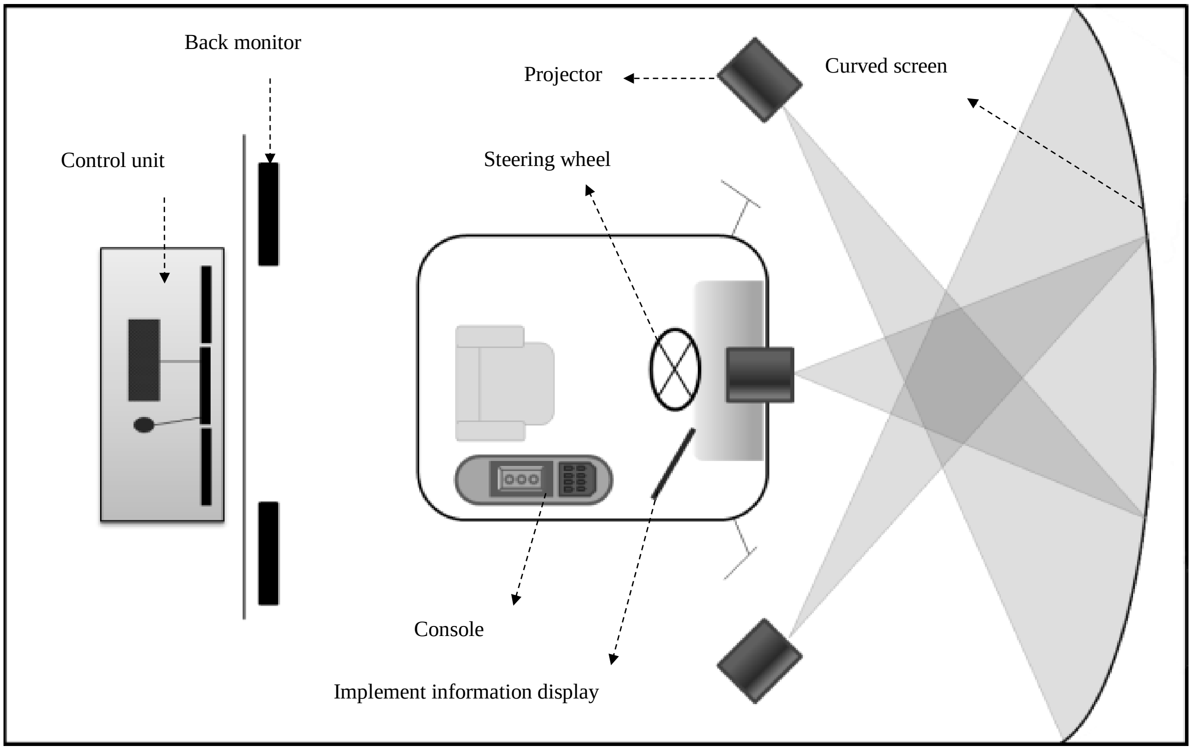
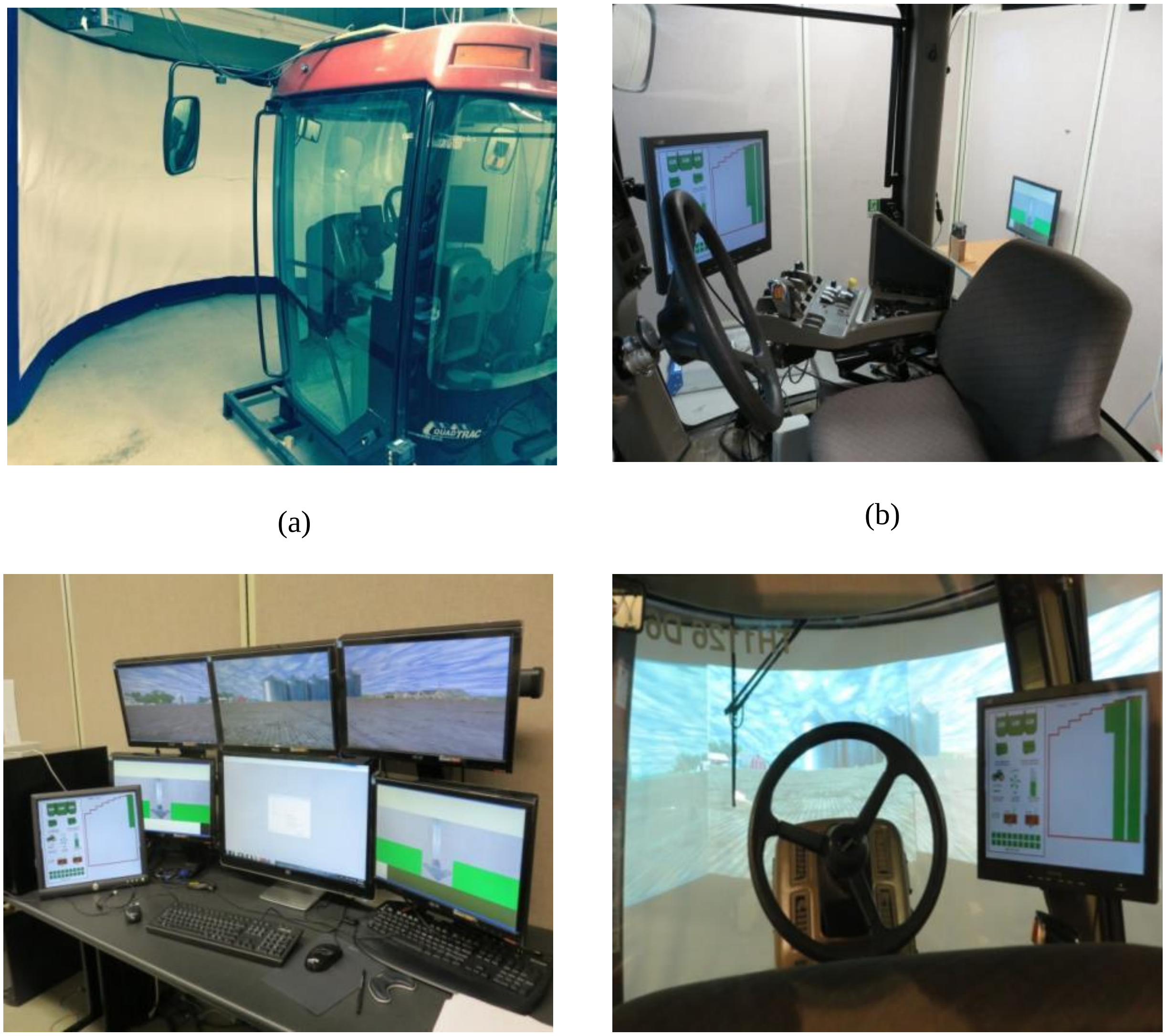


Figure 5-2. Plan view depicting the layout of the simulator

inside the simulator cab) to maintain them within specified ranges. They were also expected to monitor air seeder units that were simulated using two monitors located behind the trator cab. Consequently, simulator “drivers” were required to monitor and control the air seeder parameters as they were performing the driving function. They could use knobs, buttons and levers on the console in their right hand side to adjust the parameters that exceded the acceptable working range.

HR data were recorded using the Polar S810 heart rate monitor (Polar Electro, Finland), shown in Figure 5-4. HRV data recordings were made over driving blocks. At the end of each trial, data were transferred to a computer using a polar infrared interface. Polar precision performance software was used to analyze the data. Time and frequency-



(a) (b)

(c) (d)

Figure 5-3. The tractor driving simulator components: a) The tractor cab and the curved screen, b) Inside the cab, the steering wheel, the information display, console, and a monitor behind the cab for simulating a unit of seeder, c) The control unit, d) Visual scenery.

domain parameters were calculated over 10 min of the driving blocks. Data from the first 90s of driving blocks were ignored since there were no errors programmed to emerge at this period. Another reason was to minimize the impact of prior physical activity (i.e., moving inside the simulator cab after a break) on data.

Figure 5-4. The heart rate monitor and its components

5.4 Defining independent variables

According to the primary tasks involved in tractor-air seeder operation, two independent variables were considered in the study. Vehicle steering task automation (VSTA) was the first independent variable. It involved two levels: i) manual steering (i.e., no automation support) and ii) automatic steering (i.e., high automation support). The second independent variable was the implement control and monitoring task automation (ICMTA) that included five levels of automation support.

VSTA included manual and automatic steering modes. In manual steering mode, operators were responsible for "steering" the simulator throughout the field. Maintaining the desired pathway, avoiding path overlaps and gaps, and avoiding hazards were part of this task. In automatic steering mode, however, an auto-steer system was engaged to perform the steering task. The automatic steering mode required a purely supervisory task while the manual steering mode required a combination of physical and supervisory tasks.

The task of monitoring and controlling the seeder parameters included a supervisory task in which system errors needed to be detected and corrected immediately after their occurrence. Historically, this would have been performed by the operator without the use of any technology. In current agricultural vehicles, assistance from technology, in the form of sensors that provide information on displays inside the cab, has made it easier for operators to detect such errors. In this study, this current condition was labeled as the manual mode despite the assistance that operators receive from technology. In addition to this manual mode, levels of automation support were selected as described in the following paragraph.

For defining levels of automation for the ICMT operation, there were two options: i) to automate individual parameters of the air seeder and ii) to assume similar levels of automation for all of the parameters. Results from the task analysis showed a common trait among all of the subtasks of the air seeder control and monitoring task; all of these parameters require supervision by the operator. Whenever an action is required, it is to be completed by the operator. According to Calhoun et al. (2011), having the level of automation similar across closely coupled tasks reduces mode awareness problems. Based on this finding, the second option (i.e., to assume similar levels of automation for all of the parameters) seemed to be the better choice. Therefore, the levels of automation were assumed similar across all of the air seeder parameters.

The Parasuraman et al. (2000) model was used to define four levels of automation support. Based on this model, the ICMT task was decomposed to information processing functions of i) information acquisition, ii) information analysis, iii) decision and action selection, and iv) action implementation. Automation support was applied to each of these functions. Automation supports for these functions were selected based on the theoretical definitions presented in Chapter Two.

For information acquisition support, the computer was responsible for detecting errors and highlighting them on the display, leaving the remainder of the task to the operator. In the case of information analysis support, the computer analyzed the data and made predictions. A prediction of an error was shown in the form of a message in the message-box of the air seeder display. Operators were required to interpret the message as quickly as possible and perform the necessary action. Decision and action selection support was a condition in which the computer suggested the proper action requiring the operator to implement the action. Finally, the action implementation support was a condition where the computer performed all of the information processing functions and only informed the operator after execution of a task. This mode reflected the highest level of automation support in the experiment. Screenshots of the TAS information display in each ICMTA mode are shown in Figure 5-5.

Participants were informed that even if the steering task or ICMT was highly automated they were still responsible for supervising the system and reacting to unpredicted emergencies. For example, if there were any hazards on the path that needed to be avoided, it was the operator’s responsibility to assume control of the steering task. Similarly, if the automated system ignored any of the implement parameters that required adjustment, the operator was expected to make the adjustment. The simulator was modified to mimic both VSTA and ICMTA.

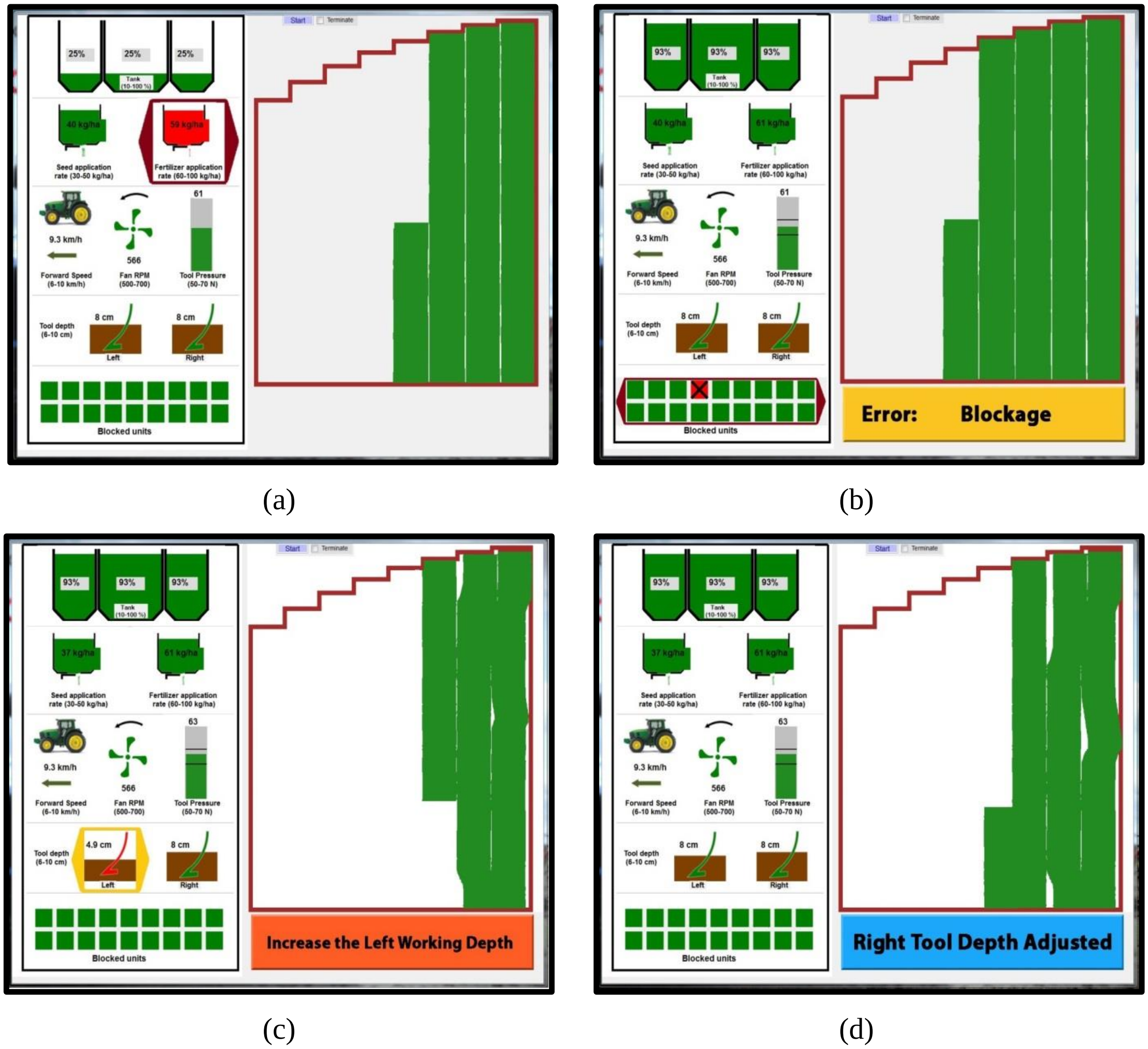


Figure 5-5. ICMTA modes: a) Information acquisition mode; a flashing box in red and yellow color appears around the parameter that needs adjustment, b) Information analysis mode; a warning message in a yellow colored box appears on the display indicating the parameter that would need adjustment, c) Decision and action selection mode; the warning message in an orange colored box offers a solution for faster removal of the error, d) Action implementation mode; the machine eliminates errors and notifies operators in a blue colored box.

5.5 Experimental design and analysis

A 2 (VSTA) × 5 (ICMTA) design was used in the study. VSTA was applied as a between subject design meaning that half of the participants performed the trial in manual steering mode and the other half in automatic steering mode. On the other hand, the

ICMTA was assumed as a within subject design so each participant was required to perform the trial in all of the automation support modes. Arranging the experiments in the form of repeated 5×5 Latin squares made it possible to avoid the learning effect and to accommodate the limited number of participants. The experimental design included six 5×5 Latin squares, sharing same columns (driving period), with participants in rows. According to the experimental condition, participants and driving periods were assumed as blocking factors, acting as random effects.

Appropriate descriptive statistics were calculated for the parameters of the dependent variables. The analyses were performed with linear mixed models using the PROC MIXED procedure in the statistical software package SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Shapiro-Wilk's test was used for normality tests prior to analysis. Data with substantial deviations were normalized by means of logarithmic or square-root transformations. Post-hoc differences of least squares means were used to determine the source of any significant effects. Statistically significant differences were accepted at the 95% of confidence level ($p < .05$) or greater. All values for parametric analyses are presented as Mean ± Standard Error.

5.6 Dependent variables

Dependent variables of this study were i) workload and ii) situation awareness of the operators. Depending on the measurement method, workload and situation awareness can be assessed either simultaneously or separately. In this study, a method for simultaneous assessment of both variables was used in order to find potential correlations between workload and situation awareness in agricultural vehicles.

As suggested in the literature, several techniques were considered for the assessment of workload, including performance and physiological measures as well as subjective workload assessment. Physiological and performance measures were taken simultaneously as the participants were performing the driving task, and the subjective evaluation tool was applied at the end of each driving session.

The DALI was used for subjective assessment of the mental workload of drivers in this experiment. In this post-trial method, drivers rated the workload experienced after completing each driving block. DALI is a promising method for measuring mental workload of drivers in the driving domain. Depending on the driving or test condition, up to seven mental workload factors can be measured using DALI: global attention demand, visual demand, auditory demand, tactile demand, stress, temporal demand, and interference. If any of these factors is not applicable, it can be simply eliminated from the questionnaire. For example, for a trial without any vibration, tactile demand has no meaning and it can be removed. According to this experimental condition, two factors (i.e., auditory demand and tactile demand) were eliminated from the DALI questionnaires. Participants were asked to rate their score on an interval scale that ranged from low (1) to high (20). A global mental workload score was calculated using the scores of the five parameters included in the questionnaire.

Primary task performance was the second measure of workload. Reaction time and accuracy of actions indicated the performance of the drivers. In this study, reaction time was defined as the period of time between the emergence of an error and the time that the driver started to make an adjustment. In terms of the accuracy of actions, three conditions were specified as failures: i) if a parameter was adjusted at a wrong time (WT), ii) if a parameter was adjusted in a wrong direction (WD), and iii) if a wrong parameter was adjusted or the parameter was ignored (WP). As an example, if the seed application rate needed to be increased, but the operator mistakenly decreased it, it would be considered a WD error. In this example, if the participant had mistakenly adjusted the fertilizer application rate instead of the seed application rate, a WP error would have been made. It should be noted that not all of the failure scenarios were applicable for all of the parameters. For instance, in the case of blockage, removing the blockage was the only available option so adjusting in a wrong direction would not be applied for this parameter.

The tractor-driving simulator was able to record real-time course of action of the drivers. Data from the simulator were reviewed separately for each participant in order to calculate performance parameters. Data from both performance parameters needed transformation to enable statistical analysis, so a logarithmic transformation was used to correct the skewness.

Heart rate variability (HRV) was the physiological measure of workload. Various HR and HRV parameters were considered in this study. Time domain parameters involved were: 1) number of heartbeats (bpm), 2) minimum RR Interval (ms), 3) average RR Interval (ms), 4) maximum RR Interval (ms), 5) standard deviation (ms), 6) max/min ratio, 7) RMSSD (ms), and 8) pNN50 (%). RMSSD is the root-mean square of differences of successive RR intervals. pNN50 is the percent of differences of adjacent RR intervals greater than 50 ms.

Parameters involved in the frequency domain were: 1) total power, TP (0.00 - 0.40 Hz), a short-term estimate of the total power of spectral density (only for the main experiment), 2) very low frequency, VLF (0.00 - 0.07 Hz) which indicates overall activity of various slow mechanisms of sympathetic function, 3) low frequency, LF (0.07 - 0.14 Hz), that reflects both sympathetic and parasympathetic activity, 4) high frequency, HF (0.14 - 0.40 Hz) that reflects parasympathetic activity, 5) LF/HF ratio that indicates overall balance between sympathetic and parasympathetic systems, 6) ISNS, normalized power of LF, and 7) IPNS, normalized power of HF.

For HRV measurement, the heart rate monitor, described in section 5.3, was used. The HRV data were recorded during the test for later analysis. Most of the parameters required transformation to enable statistical analysis. Depending on the data values, logarithmic or square-root transformations were used to normalize data and correct the skewness in distribution of variance of parameters.

SART was used as the subjective measure of situation awareness. This rating scale measured 10 dimensions of situation awareness of drivers. These dimensions were categorized in three distinctive groups: i) demand on attentional resources, ii) supply of attentional resources and iii) understanding. Demand on attentional resources included questions regarding instability, complexity and variability of a situation. Supply of attentional resources reflected arousal, concentration, division and spare mental resources during a driving condition. Finally, understanding of a situation depended on the quality and quantity of the information provided as well as familiarity with the situation. A value for each of these categories was derived by taking the average of responses to questions included in the appropriate categories. Lastly, the combined rate for situation awareness was inferred by subtracting the average score of demand from the sum of average scores of understanding and supply. Participants completed the SART questionnaire immediately at the end each driving block, prior to completing the DALI questionnaire.

In addition to the measures described above, a short biographic questionnaire was given to each participant to ensure that subjects met the qualifications of the study and to record their background information. The survey included queries regarding the subject's gender, age, car and tractor driving experience, and previous experience with the TDS. Finally, a post-trial questionnaire was used to get subjective feedback on the simulator experience and the test procedure. The questionnare was comprised of four queries. As the first query, participants were asked to report the percentage of time that they spent viewing different items while driving the simulator. The items included map/navigation display, the air seeder information display, visual scenery (i.e., working path), the monitors behind cab used for air seeder simulation, and finally any other items. For the second question, the subjects described the ease with which they were able to find the appropriate information on the instrument panel inside the simulator's cab. For this query, they could choose from 5 difficulty levels: easy, very easy, neutral, difficult, and very difficult . The reamining were two open-ended questions to which the participants could voluntarily respond. One asked participants to reflect on the experimental protocol of answering identical questionnaires of SART and DALI after each driving period. The last question asked about any compliments, suggestions or complaints of the participants on the experiment, display configuration, or the simulator.

5.7 Procedure

For all of the tests, subjects received explanations of the test procedure. They were provided with necessary instructions upon arrival to the Lab. They were required to complete a biographic questionnaire and read and sign a consent letter (Appendix B). A 15 min training session was administered in order to make subjects comfortable with the test procedure and to allow them to familiarize themselves with the simulator and the implement control console. The driving task in the training session included short periods (2-3 min) of all ICMTA modes. After the training session, they completed the main trial that included five driving blocks, 12 min each. At the end of each driving block, paper-based queries of DALI and SART were given to the subjects. At the end of the last driving block, subjects were also required to complete a post-trial questionnaire. Short breaks (maximum 10 min) were given as needed following the training session and each of the experimental blocks. The entire experimental session lasted between 2 and 2.5 h. All of the forms that were used in the experiment are shown in Appendix B.

5.8 Hypotheses

It was hypothesized that ICMTA levels and VSTA would affect the operator's mental workload, reaction time, accuracy of actions, and HRV. With respect to mental workload, lower scores were expected by increasing ICMTA and VSTA levels. Similarly, shorter reaction time and number of errors for the highly automated condition were expected. This is based on the assumption that the automation of information acquisition, information analysis, and decision and action selection provides added information and performance enhancing features for the task. Moreover, automation of action implementation greatly reduces task physical activity requirement. With respect to HRV, it was expected that the 0.1 Hz component, as the widely used representative of HRV, would increase as the level of ICMTA and VSTA increased. This was based on the assumption that the increase in automation level would decrease mental workload. The other HRV parameters were expected to show variations to the change in automation condition as the level of mental workload and physical activity would be altered in the experiments.

With respect to situation awareness, it was hypothesized that ICMTA and VSTA increase would result in lower scores. This is because of the out-of-the-loop effect. As the automation is applied to higher levels of information processing functions (i.e., from information acquisition to action implementation), it was anticipated that operators would experience lower involvment in the task loop.

Chapter 6

RESULTS: PRELIMINARY EXPERIMENT

6.1 Introduction

This chapter reports on the results of the pilot study. Each sub-section of this chapter presents results from one measurement technique. At the end of this chapter, a section describes the lessons that were learnt from the pilot study and the changes that were made for conducting the main experiment.

6.2 Mental workload

Figure 6-1 shows the means of DALI parameters for the preliminary experiment in manual and automatic steering modes. In all of the cases, manual steering resulted in higher global mental workload and higher values for its components, however, this was not identified as significant effect. Means of DALI parameters in different ICMTA levels are shown in Figure 6-2. ICMTA effect was found on all of the parameters with the exception of interference. The VSTA × ICMTA interaction effect was only found on attention demand.

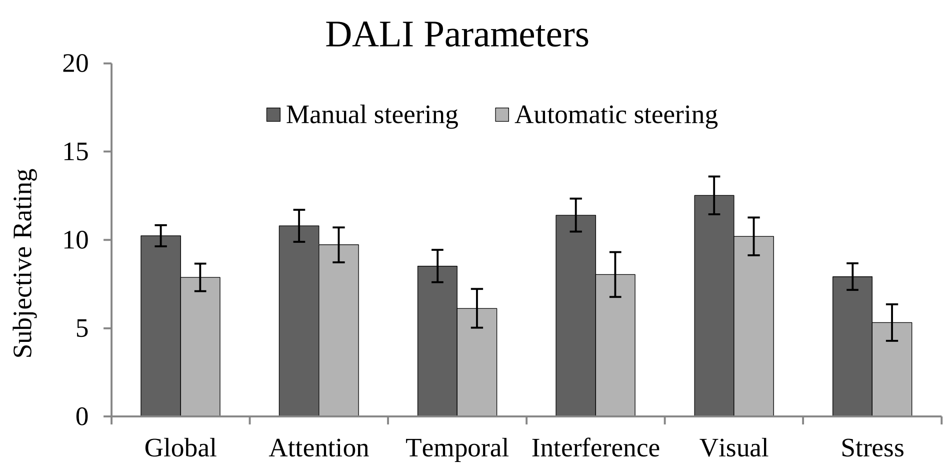


Figure 6-1. Mental workload and its parameters for different VSTA levels from the preliminary experiment.

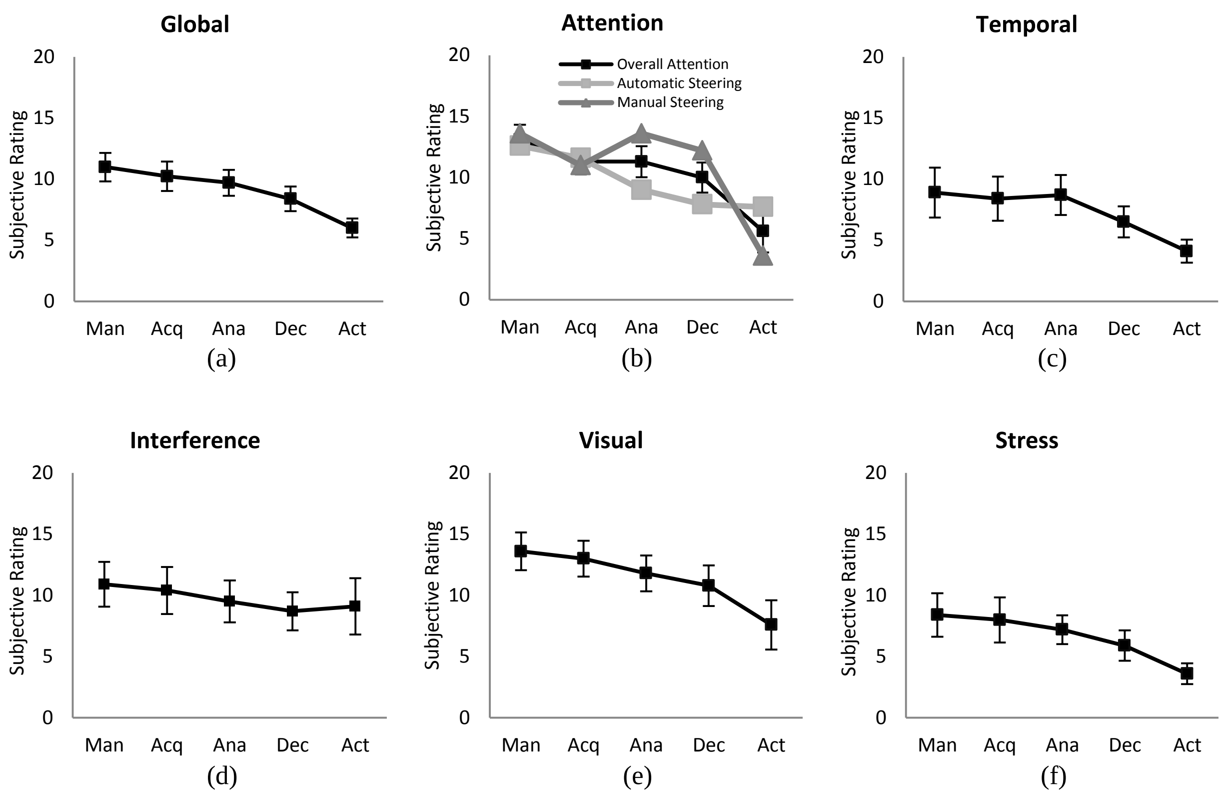


Figure 6-2. Mental workload parameters for different ICMTA modes: Manual (Man), Information Acquisition (Acq), Information Analysis (Ana), Decision and action Selection (Dec), and Action Implementation (Act).

6.2.1 Global workload

Looking at the *global workload* results (Figure 6-2a), a decreasing trend was observed with increasing automation level. ANOVA showed significant differences among ICMTA levels in terms of subjective assessment of workload by the operators, F (4, 28) = 12.73, p < .001, ω^2_p = 0.484. Table 6-1 shows pairwise comparison of the ICMTA levels. The first three conditions ,i.e., Manual (M =11.0, SE = 1.1), Information Acquisition (M =10.2, SE = 1.2), and Information Analysis (M = 9.7, SE = 1.1) modes, did not result in any significant differences with one another. Having the lowest mean value, the Action Implementation mode (M = 6.0, SE = 0.8) showed significantly different effect from the other ICMTA levels. Decision and Action Selection mode (M = 8.4, SE = 1.0) also had a different effect on mental workload compared to the Manual and Information Acquisition modes. There was no effect of either VSTA or VSTA × ICMTA interaction on global workload, $p > .05$.

Table 6-1. Multiple comparison table on global mental workload of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.8	0.8	28	0.98	0.335	-2.35	0.83
	Ana	1.3	0.8	28	1.65	0.110	-2.87	0.31
	Dec	2.6	0.8	28	3.36	0.002	-4.19	-1.01
	Act	5.0	0.8	28	6.43	<.001	-6.57	-3.39
Acq	Ana	0.5	0.8	28	0.67	0.507	-1.07	2.11
	Dec	1.8	0.8	28	2.38	0.025	0.25	3.43
	Act	4.2	0.8	28	5.45	<.001	2.63	5.81
Ana	Dec	1.3	0.8	28	1.70	0.099	-0.27	2.91
	Act	3.7	0.8	28	4.78	<.001	-5.29	-2.11
Dec	Act	2.4	0.8	28	3.07	0.005	-3.97	-0.79

6.2.2 Attentional demand

Attention is used in the DALI questionnaire to evaluate the attention – to think about, to decide, to choose, to look for – required by the task or activity. ANOVA showed no main effect of VSTA on this parameter, $p > .05$. There was a significant main effect of ICMTA, F (4, 28) = 17.5, $p < .001$, $\omega^2_p = .570$. For attentional demand (Figure 6-2b), similar to global workload results, no differences were found among Manual (M = 13.1, SE = 1.2), Information Acquisition (M = 11.3, SE = 1.0), and Information Analysis (M = 11.3, SE = 1.3) modes (Table 6-2). Decision and Action Selection mode (M = 10.0, SE = 1.2) showed significant differences with Manual and Action Implementation (M = 5.6, SE = 1.7) modes. By significantly reducing the attentional demand, Action Implementation mode had different effect compared to the rest of the ICMTA levels.

There was a significant VSTA × ICMTA interaction effect on *temporal demand*, F (4, 28) = 7.11, $p < .001$, $\omega^2_p = .329$. Manual modes in different VSTA modes did not show any differences from each other, $p > 0.05$. A similar condition was observed in the cases of Information Acquisition and Action Implementation modes, however, differences were found for the two remaining levels of ICMTA. Manual Steering resulted in significantly higher scores for Information Analysis, and Decision and Action Selection modes compared to Automatic Steering, $p < 0.05$.

Table 6-2. Multiple comparison table on Attentional demand of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	1.8	1.0	28	1.88	0.070	-3.76	0.16
	Ana	1.8	1.0	28	1.88	0.070	-3.76	0.16
	Dec	3.1	1.0	28	3.24	0.003	-5.06	-1.14
	Act	7.5	1.0	28	7.84	<.001	-9.46	-5.54
Acq	Ana	0.0	1.0	28	0.00	1.000	-1.96	1.96
	Dec	1.3	1.0	28	1.36	0.185	-0.66	3.26
	Act	5.7	1.0	28	5.96	<.001	3.74	7.66
Ana	Dec	1.3	1.0	28	1.36	0.185	-0.66	3.26
	Act	5.7	1.0	28	5.96	<.001	-7.66	-3.74
Dec	Act	4.4	1.0	28	4.60	<.001	-6.36	-2.44

6.2.3 Temporal demand

Temporal demand shows the time pressure associated with the whole activity. There was a main effect of ICMTA on *temporal demand*, F (4, 28) = 4.53, p <.01, ω^2_p = .217, but no significant effects of VSTA or VSTA × ICMTA automation interaction were observed, p > .05. Figure 6-2c demonstrates results of attentional demand in different ICMTA levels. According to the table of pairwise analysis (Table 6-3), Manual (M = 8.9, SE = 2.0), Information Acquisition (M= 8.4, SE = 1.8) and Information Analysis (M = 8.7, SE = 1.6) modes had similar effects on the temporal demand of the operators. Action Implementation mode caused the lowest temporal demand with a mean of 4.1 (SE = 0.9). This was not significantly different from the effect of Decision and Action Selection (M = 6.5, SE = 1.2), but was significantly different from the other automation types.

Table 6-3. Multiple comparison table on temporal demand of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.5	1.4	28	0.37	0.715	-3.27	2.27
	Ana	0.2	1.4	28	0.15	0.884	-2.97	2.57
	Dec	2.4	1.4	28	1.77	0.087	-5.17	0.37
	Act	4.8	1.4	28	3.55	0.001	-7.57	-2.03
Acq	Ana	0.3	1.4	28	0.22	0.826	-3.07	2.47
	Dec	1.9	1.4	28	1.40	0.171	-0.87	4.67
	Act	4.3	1.4	28	3.18	0.004	1.53	7.07
Ana	Dec	2.2	1.4	28	1.63	0.115	-0.57	4.97
	Act	4.6	1.4	28	3.40	0.002	-7.37	-1.83
Dec	Act	2.4	1.4	28	1.77	0.087	-5.17	0.37

6.2.4 Visual demand

In the case of *visual demand* that was required during the test to perform the activity, ICMTA effect was significant, F (4, 28) = 5.33, p <.01, ω^2_p = .239. Action Implementation mode (M = 7.6, SE = 2.0) resulted in significantly lower subjective ratings (Figure 6-2e). The remainder of the ICMTA levels resulted in moderate to above moderate values (Manual (M =13.6, SE = 1.6), Information Acquisition (M =13.0, SE = 1.5), and Information Analysis (M = 11.8, SE = 1.5), Decision and Action Selection (M = 10.8, SE = 1.7)), showing no differences from one another (Table 6-4). ANOVA indicated no main effect of VSTA or VSTA × ICMTA interaction, p > .05.

Table 6-4. Multiple comparison table on visual demand of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.6	1.4	28	0.41	0.682	-3.57	2.37
	Ana	1.8	1.4	28	1.24	0.224	-4.77	1.17
	Dec	2.8	1.4	28	1.93	0.063	-5.77	0.17
	Act	6.0	1.4	28	4.14	0.000	-8.97	-3.03
Acq	Ana	1.2	1.4	28	0.83	0.415	-1.77	4.17
	Dec	2.2	1.4	28	1.52	0.140	-0.77	5.17
	Act	5.4	1.4	28	3.73	0.001	2.43	8.37
Ana	Dec	1.0	1.4	28	0.69	0.496	-1.97	3.97
	Act	4.2	1.4	28	2.90	0.007	-7.17	-1.23
Dec	Act	3.2	1.4	28	2.21	0.036	-6.17	-0.23

6.2.5 Situational stress

All of the automation types imposed a lower than moderate level of *stress* on operators (Figure 6-2f). Similar to visual demand, there was a main effect of ICMTA on this parameter, F (4, 28) = 6.36, p <.001, ω^2_p = .300. Action Implementation mode (M = 3.6, SE = 0.8) triggered the lowest scores, being the least stressful situation. A significant difference was also found between Manual (M = 8.4, SE = 1.8) and Decision and Action Selection (M = 5.9, SE = 1.2) modes, as shown in table 6-5. Information Acquisition (M =8.0, SE = 1.8), and Information Analysis (M = 7.2, SE = 1.2) caused similar levels of stress. no main effect of VSTA or VSTA × ICMTA interaction were observed, p > .05.

Table 6-5. Multiple comparison table on situational stress of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.4	1.1	28	0.37	0.716	-2.63	1.83
	Ana	1.2	1.1	28	1.1	0.279	-3.43	1.03
	Dec	2.5	1.1	28	2.3	0.029	-4.73	-0.27
	Act	4.8	1.1	28	4.41	0.000	-7.03	-2.57
Acq	Ana	0.8	1.1	28	0.74	0.468	-1.43	3.03
	Dec	2.1	1.1	28	1.93	0.064	-0.13	4.33
	Act	4.4	1.1	28	4.05	0.000	2.17	6.63
Ana	Dec	3.6	1.1	28	3.31	0.003	-5.83	-1.37
	Act	1.3	1.1	28	1.2	0.242	-0.93	3.53
Dec	Act	2.3	1.1	28	2.11	0.044	-4.53	-0.07

6.3 Performance

Results from the preliminary experiment showed that driving with the auto-steer system (M = 2.84, SE = 0.42 s) caused higher average reaction time compared to the manual steering mode (M = 2.09, SE = 0.42 s). By contrast, the number of failures was higher in manual steering mode (27 vs. 44). Operators made 71 errors. The number of errors concerning wrong timing (WT) was 47, for adjusting in the wrong direction (WD) was 20, and for failing to detect the need for parameter adjustment (WP) was 4.

Figure 6-3 shows means of reaction time and human errors for different VSTA and ICMTA levels. Despite the lower values in automatic steering mode, the statistical analysis showed no VSTA effect on reaction time or number of errors, $p > .05$, nor was the VSTA × ICMTA interaction effect, $p > 0.5$.

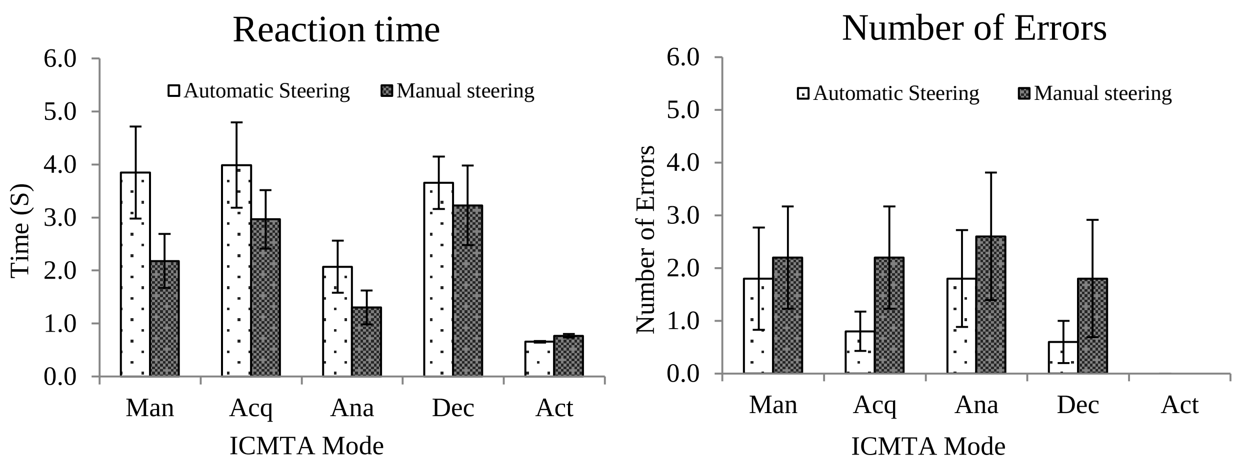


Figure 6-3. Reaction time and number of errors for different VSTA and ICMTA levels for preliminary experiment.

ICMTA, on the other hand, displayed significant effects on reaction time, F (4, 28) = 17.63, p <.001, ω^2_p = .571, and number of errors, F (4, 28) = 4.90, p <.01, ω^2_p = .238. It was hypothesized that as the level of ICMTA was increased, reaction time would decrease. The observed results (Table 6-6), however, showed inconsistencies; no differences were found among Manual (M = 3.01, SE = 0.55 s), Information Acquisition (M = 3.48, SE = 0.50s) and Decision and Action Selection (M = 3.44, SE = 0.43 s) modes. Post-hoc analysis indicated that Information Analysis (M = 1.69, SE = 0.30 s) and Action Implementation (M = 0.71, SE = 0.03 s) supports significantly reduced the reaction time. In terms of accuracy of actions (Table 6-7), similar numbers of errors were made in Manual (M = 2.00, SE = 0.65), Information Acquisition (M = 1.50, SE = 0.54), Information Analysis (M = 2.2, SE = 0.73) and Decision and Action Selection (M = 1.2, SE = 0.60) modes, unlike the Action Implementation mode with no errors.

Table 6-6. Multiple comparison table on reaction time of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.5	0.4	28	1.12	0.270	-0.38	1.30
	Ana	1.3	0.4	28	3.23	0.003	-2.17	-0.49
	Dec	0.4	0.4	28	1.04	0.307	-0.41	1.27
	Act	2.3	0.4	28	5.6	<.001	-3.14	-1.46
Acq	Ana	1.8	0.4	28	4.36	0.000	0.95	2.63
	Dec	0.0	0.4	28	0.09	0.933	-0.81	0.88
	Act	2.8	0.4	28	6.72	<.001	1.92	3.60
Ana	Dec	1.8	0.4	28	4.27	<.001	-2.60	-0.91
	Act	1.0	0.4	28	2.37	0.025	-1.81	-0.13
Dec	Act	2.7	0.4	28	6.64	<.001	-3.57	-1.89

Table 6-7. Multiple comparison table on number of errors of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.5	0.6	28	0.90	0.374	-1.63	0.63
	Ana	0.2	0.6	28	0.36	0.721	-0.93	1.33
	Dec	0.8	0.6	28	1.44	0.160	-1.93	0.33
	Act	2.0	0.6	28	3.61	0.001	-3.13	-0.87
Acq	Ana	0.7	0.6	28	1.26	0.217	-1.83	0.43
	Dec	0.3	0.6	28	0.54	0.592	-0.83	1.43
	Act	1.5	0.6	28	2.71	0.011	0.37	2.63
Ana	Dec	1.0	0.6	28	1.81	0.082	-0.13	2.13
	Act	2.2	0.6	28	3.97	0.001	-3.33	-1.07
Dec	Act	1.2	0.6	28	2.17	0.039	-2.33	-0.07

6.4 HRV

The results from the preliminary experiment did not show any significant effect of VSTA or ICMTA on HRV parameters of participants, $p > .05$. This result could be due to two reasons: i) the insensitivity of this measure in the tractor-driving simulator, and ii) the experimental condition. It has been stated that HRV is affected by age (Jensen-Urstad et al. 1997; Takahashi et al. 2012; Thayer et al. 2009), to a lesser degree by gender (Jensen-Urstad et al. 1997) genetic composition, and environment (Taelman et al. 2008). Age of subjects, in this study, varied from 20 to 35 which is a wide range in the case of

HRV. Furthermore, the subjects were from different ethnicities; two Caucasian, two Hispanic, five Asian, and one African participated in the experiment. Another factor that affects HRV is the gender of subjects. In the experiment, there was one female participant and nine male participants (i.e., 10% of the results came from a different gender). Participants also differed in terms of car and tractor driving experience. Car driving experience of the participants ranged from less than 1 year to 15 years. Five participants did not have any tractor driving experience, while two of the participants had over 10 years of tractor driving experience.

It should be pointed out that there are some controversial findings on the influence of gender or ethinicity on HRV. Some studies reported greater HRV in female subjects (Kim & Woo 2011; Ryan et al. 1994), while others reported the converse (Ramaekers et al. 1998). Wang et al. (2005) found that the relative contributions of genetic and environmental factors to HRV parameters in African-American youth were similar to those in European American youth and the same in male and female subjects.

6.5 Situation Awareness

Following statistical analysis of the data, scores of SART-combined and its three components were studied. Component ratings ranged from 1 to 20. SART-combined was calculated from the means of its three components. As shown in Figure 6-4, the auto-steer mode resulted in slightly higher situation awareness of drivers. Both VSTA levels imposed almost equal demand on attentional resources and understanding, although automatic steering allowed higher values for supplying of attentional resources. ANOVA revealed no VSTA, ICMTA or VSTA × ICMTA interaction effect on SART-combined, $p > .05$.

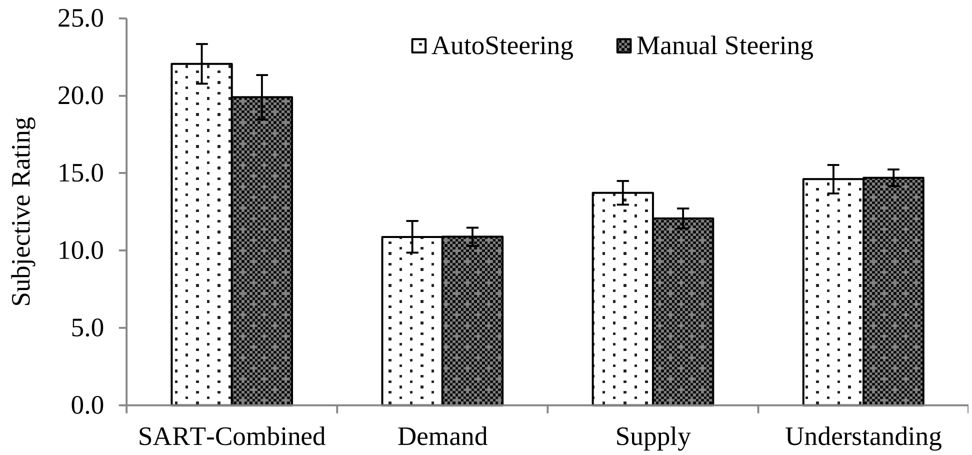


Figure 6-4. Situation awareness of operators and its components for different VSTA levels in preliminary experiment.

For the three dimensions of SART, no VSTA or VSTA × ICMTA interaction effects were found, p > .05. An ICMTA effect was found on two dimensions: i) demand on attentional resources, F (4, 28) = 4.81, p <.01, ω^2_p = .229, and ii) supply of attentional resources, F (4, 28) = 6.12, p <.01, ω^2_p = .290. According to the post-hoc differences of least square means, in the case of demand on attentional resources (Table 6-8), for all of the observations, Action Implementation (M = 8.7, SE = 1.5) had a different effect compared to the rest of ICMTA levels (Manual (M =11.4, SE = 1.3), Information

Table 6-8. Multiple comparison table on demand on attentional resources of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.2	0.8	28	0.19	0.852	-1.78	1.48
	Ana	0.3	0.8	28	0.33	0.746	-1.37	1.89
	Dec	0.1	0.8	28	0.15	0.881	-1.51	1.75
	Act	2.7	0.8	28	3.37	0.002	-4.31	-1.05
Acq	Ana	0.4	0.8	28	0.52	0.610	-2.04	1.22
	Dec	0.3	0.8	28	0.34	0.737	-1.90	1.36
	Act	2.5	0.8	28	3.18	0.004	0.90	4.16
Ana	Dec	0.1	0.8	28	0.18	0.862	-1.49	1.77
	Act	2.9	0.8	28	3.70	0.001	-4.57	-1.31
Dec	Act	2.8	0.8	28	3.52	0.002	-4.43	-1.17

Acquisition (M =11.2, SE = 1.2), and Information Analysis (M = 11.6, SE = 1.3), Decision and Action Selection (M = 11.5, SE = 1.3)). Based on the means of demand (Figure 6-5), Action Implementation reduced demand on attentional resources by 13.5% compared to the other ICMTA levels.

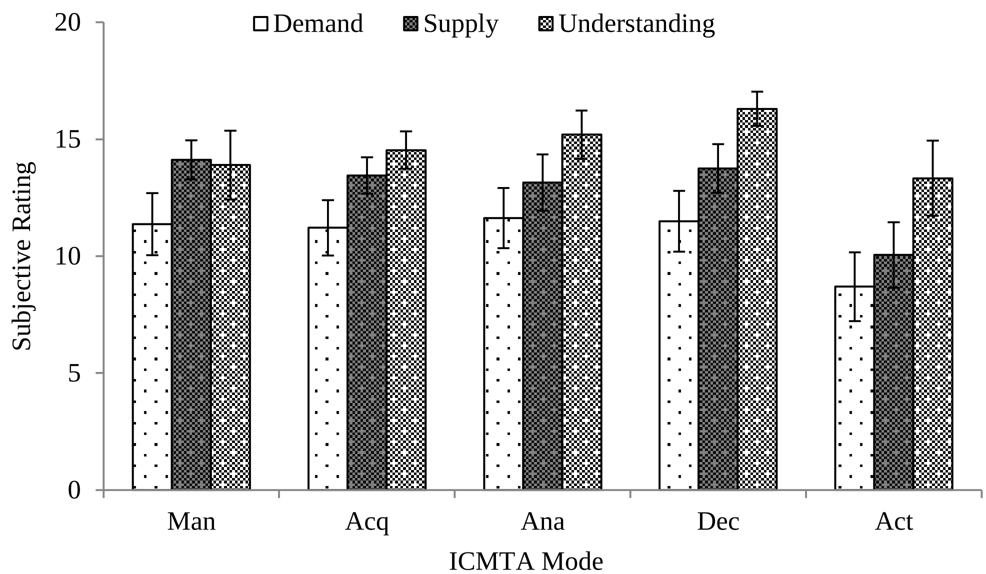


Figure 6-5. Situation awareness components for five ICMTA conditions.

Similar results were found in the case of supply of attentional resources where the Action Implementation mode (M = 10.1, SE = 1.4) resulted in 17.5% smaller values, showing significant differences with the rest of ICMTA levels (Table 6-9). Manual (M = 14.2, SE = 0.8), Information Acquisition (M = 13.5, SE = 0.8), and Information Analysis (M = 13.2, SE = 1.2), Decision and Action Selection (M = 13.8, SE = 1.0) resulted in similar values.

Based on the SART-combined formula, reduction in demand would be a desirable outcome while reduction in supply of attentional resources would not. Overall, the Action Implementation mode appears to not be the best choice if the situation awareness of the

Table 6-9. Multiple comparison table on supply of attentional resources of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.7	0.9	28	0.70	0.487	-2.58	1.26
	Ana	1.0	0.9	28	1.05	0.305	-2.90	0.94
	Dec	0.4	0.9	28	0.41	0.688	-2.30	1.54
	Act	4.1	0.9	28	4.35	0.000	-6.00	-2.16
Acq	Ana	0.3	0.9	28	0.34	0.735	-1.60	2.24
	Dec	0.3	0.9	28	0.30	0.767	-2.20	1.64
	Act	3.4	0.9	28	3.65	0.001	1.50	5.34
Ana	Dec	0.6	0.9	28	0.64	0.527	-2.52	1.32
	Act	3.1	0.9	28	3.31	0.003	-5.02	-1.18
Dec	Act	3.7	0.9	28	3.95	0.001	-5.62	-1.78

drivers needs to remain at high levels. Therefore, the driver needs to be involved in the task loop to some extent.

6.6 Correlations

Simple correlation analyses were conducted using the Pearson coefficient in order to identify any significant relationships among the variables with significant results. The following approximation (Table 6-10) suggested by Cohen (1988) for the Behavioral Sciences was used to interpret the strength of positive or negative correlations.

Table 6-10. Correlation strength approximation (Cohen 1988).

Correlation coefficient (r)	Interpretation
$0.1 \leq r < 0.3$	Small effect
$0.3 \leq r < 0.5$	Medium effect
$r \geq 0.5$	Strong effect

For the preliminary experiment, HRV parameters were not included in the correlation analysis since they did not show any significant effects. Table 6-11 shows Pearson coefficients for global mental workload and situation awareness score, reaction time and number of errors. No association was observed between mental workload and situation awareness. Reaction time also did not show any correlations with mental

Table 6-11. Pearson correlations for perceived workload, situation awareness, reaction time and number of errors in preliminary experiments.

	Workload	Situation Awareness	Reaction Time
Situation Awareness	-0.001		
Reaction Time	0.22	0.26	
Number of Errors	0.48**	-0.34*	-0.13

** $p < 0.01$ level.
* $p < 0.05$ level.

workload, situation awareness and number of errors. Number of errors, however, showed positive correlation with workload and negative correlation with situation awareness.

Pearson correlation coefficients of workload and performance components are provided in Table 6-12. Attentional demand showed positive correlations with all of the components, although, the correlations with interference and reaction time were not statistically significant. A strong correlation between attentional and visual demands (r = 0.75) suggests that subjects allocated considerable attention to the implement monitoring task and the mapping system using the information display. Visual demand did not show any significant correlations with other components of workload and performance. Stress was found to have strong correlations with attention, temporal demand and interference. The strong positive correlation with temporal demand (r = 0.87) shows that higher time-

Table 6-12. Pearson correlations of workload and performance components in preliminary experiments.

	attention	Visual	Stress	Temporal	Interference	global	Reaction Time
Visual	0.75**						
Stress	0.31*	0.36					
Temporal	0.35*	0.41	0.87**				
Interference	0.07	-0.06	0.44**	0.45**			
Global	0.68**	0.68**	0.82**	0.86**	0.55**		
Reaction Time	0.27	0.07	0.23	0.21	0.04	0.22	
Number of Errors	0.36*	0.28	0.23	0.41**	0.41**	0.48**	-0.13

** $p < 0.01$ level.
* $p < 0.05$ level.

demand resulted in higher stress level in subjects. Temporal demand had positive correlations with all of the parameters, however, the correlations with visual demand and reaction time were not statistically significant. Besides the aforementioned correlations, interference showed a moderate positive correlation with number of errors. Remarkably, the reaction time did not show any significant correlations with mental workload components or number of errors. Large positive correlations of the global score of mental workload with its components occurred because this score was derived from those components. Global score of mental workload also showed a moderate positive correlation with number of errors.

Table 6-13 presents the results of Pearson correlation analysis on the components of mental workload and situation awareness. The demand component of situation awareness showed significant positive correlation with all of the components of workload and the remaining components of situation awareness. Supply and understanding components of situation awareness also showed positive correlations with the components of workload, however, they did not have significant correlations with stress, temporal demand and interference.

Table 6-13. Pearson correlations of workload and situation awareness components in preliminary experiments.

	Attention	Visual	Stress	Temporal	Interference	Demand	Supply
Demand	0.61**	0.59**	0.39**	0.53**	0.32*		
Supply	0.65**	0.51**	0.17	0.28	0.09	0.76**	
Understanding	0.36*	0.43**	0.17	0.24	0.02	0.71**	0.52**

** $p < 0.01$ level.
* $p < 0.05$ level.

Correlations among parameters of performance and situation awareness are shown in Table 6-14. Correlation analysis showed significant positive association of reaction time with supply of attentional resources. Number of errors was also positively correlated with demand and supply. It was expected that higher task demand would result in greater number of errors. Positive correlations of performance parameters with supply of attentional resources are contradictory.

Table 6-14. Pearson correlations of performance and components of situation awareness in the preliminary experiment.

	Demand	Supply	Understanding
Reaction Time	0.18	0.36*	0.18
Number of Errors	0.44**	0.30*	-0.02

** $p < 0.01$ level.
* $p < 0.05$ level.

6.7 Post-experiment questionnaire

The post-experiment questionnaire was used to collect subjective feedback on the test procedure to inform future studies in the simulator. A summary of responses of participants to quantitative and qualitative queries and open-ended questions in the post-experiment questionnaire is reported in this section. Results are provided in percentile format considering that some of the volunteers did not answer all of the queries.

First query was regarding percentage of time that participants spent viewing different items while driving the simulator. For this question, answers from one subject were ignored as he did not respond accurately (the sum was not 100%). Overal, participants allocated 43% of their time to monitor the mapping system. They also allocated 35% of their time on monitoring the implement parameters. Looking ahead of the cab (visual scenery, path) and behind the cab required 8% and 10% of the participant's time, respectively. Finally, the average time they spent viewing everything else was 4%. As shown in figure 6-6, drivers in manual steering mode spent much more time on looking at mapping system. On the other hand, the time spent on looking at air-seeder display, visual scenery, implement monitoring, etc. was longer for the participants in the automatic steering mode. This result was expected as the automatic steering mode did not require constant information acquisiton from the mapping system for navigation, so participants could allocate more time on monitoring other components.

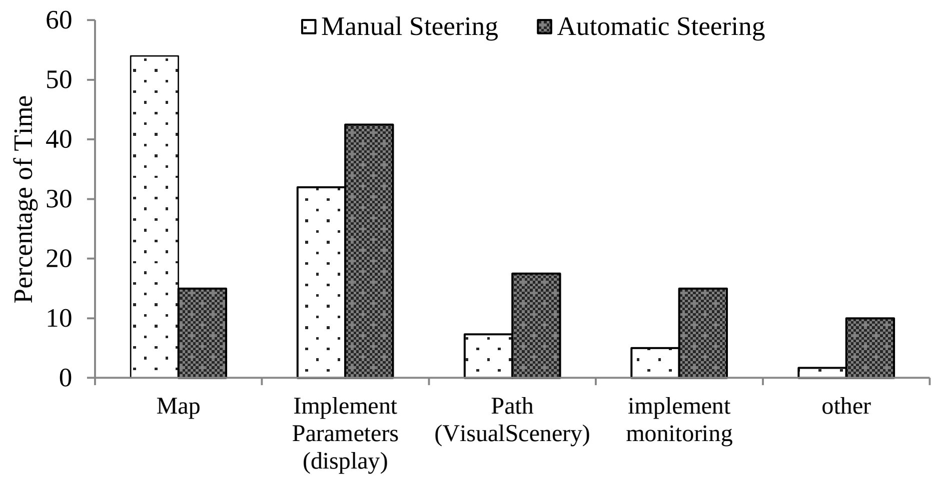


Figure 6-6. The average time spent on supervising various item in the simulator in different VSTA modes.

For the second question, the subjects described the ease with which they were able to find the appropriate parameter controler on the console inside the simulator's cab. As shown in Figure 6-7, two subjects found it very easy to find the appropriate items on the console, three found it easy, one found it neither easy nor difficult, and one subject found

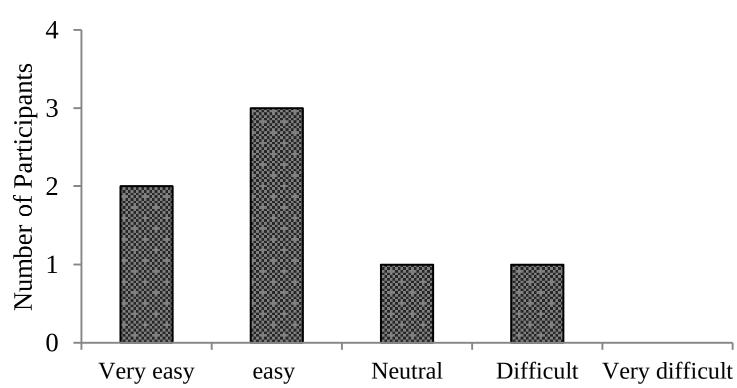


Figure 6-7. Ease of air seeder parameter finding on the console for participants.

it difficult. The rest of the participants did not rate their experience. In average, participants in the manual steering mode and automatic steering mode respectively found it very easy and easy to locate parameters on the console.

The next question asked participants to reflect on the experimental protocol of answering identical questionnaires of SART and DALI after each driving period. Positive feedback was collected from eight subjects. Examples of comments in favour of the test protocol are as follow:

"It is good because I can answer questions based on the experience I just had for each session"

"It makes sense to get those questions answered every time to assess each aspect of the research"

One subject found it tedious. Another participant expressed that: *"I don't really know if my answers are accurate because the questions are too general"*

The last question asked about any compliments, suggestions or complaints of the subjects on the experiment, display configuration, and the simulator. From the partiipants who chose to answer this query, some of the subjects expressed positive feelings regarding participation in the experiment. Comment from one participants is noted below:

"Overall, the experience was quite good for me. It added value to me in understanding the user interface from a different perspective."

Decision and Action Selection mode was identified as the favorite mode for two participants, as they expressed:

"I like the last one, when the message is telling the operator what to do to fix the error so I noticed the warning sign right away and could easily make adjustments."

"I found decision-making condition easiest to handle with the exception of the action automation mode"

Suggestions included adding extra features to the simulator including a touch screen display, configuration of buttons in order of their appearance on the screen, relocating the message box on the screen, and adding sound alerts to warning messages. Finally, the tractor noise was annoying for one subject.

All of the comments from participants to open-ended queries can be found in Appendix C.

6.8 Lessons learnt from the pilot study

The pilot study was performed with participants from a wide age range population. The individuals were also quite different in car and tractor driving experience. During the trials, it could be observed that some of participants adjusted easily to the test environment and performed the task confidently. On the other hand, some of the participants with lower experience needed more time for training. Based on this observation, it was decided that for the main experiment, only participants with tractor driving experience would be recruited.

Given that the results of HRV did not produce any significant results, it was suspected that this might be mainly due to the age range of participants. Hence, for the main experiment a narrow age limit, i.e. 18-26, was used. Considering that most of university students fitted in this age limit, difficulty of finding volunteers remained in a similar level as for pilot study. Based on the findings in the literature, as reported in section 6-6, ethnicity and gender of participants were not considred as restraining factors for the main experiment. For the main experiment, 28 Caucasian and 2 participants from a different ethnicity were volunteered. This can be a very good indicative of the population of operators of farm equipment in Canada.

Chapter 7

RESULTS: MAIN EXPERIMENT

7.1 Introduction

Following the pilot study, a full-scale experiment with 30 participants was performed to examine the research hypotheses with higher statistical power. The results from the full-scale experiment are reported in this chapter. Similar to the previous chapter, each sub-section of this chapter is concerning one independent variable. Furthermore, results of correlation analyses are also provided.

7.2 Mental workload

The subjective feedback revealed moderate levels of experienced mental workload for most of the driving conditions in the main experiment. Figure 7-1 shows means of DALI parameters in different VSTA modes. In most cases, parameters were reported at moderate levels. The means of all parameters in automatic steering mode were smaller than the means of parameters in manual steering mode, but these differences were not statistically significant. Means of DALI parameters in different ICMTA levels are shown in Figure 7-2.

7.2.1 Global workload

There was no effect of either VSTA or VSTA × ICMTA interaction on global workload, $p > .05$. On the other hand, there was a main effect of ICMTA on global workload, $F(4, 102) = 28.91$, $p < .001$, $\omega^2 p = 0.439$, so changing levels of ICMTA resulted in some significant differences among subjective feedback on global workload

(Figure 7-2a). The differences were mostly due to the lower mental workload in the Action Implementation

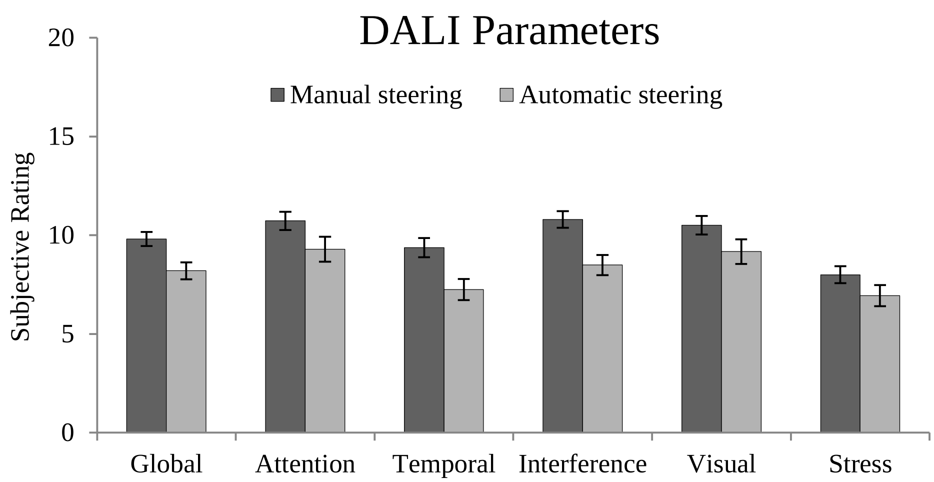


Figure 7-1. Mental workload components for different VSTA levels from the main experiment.

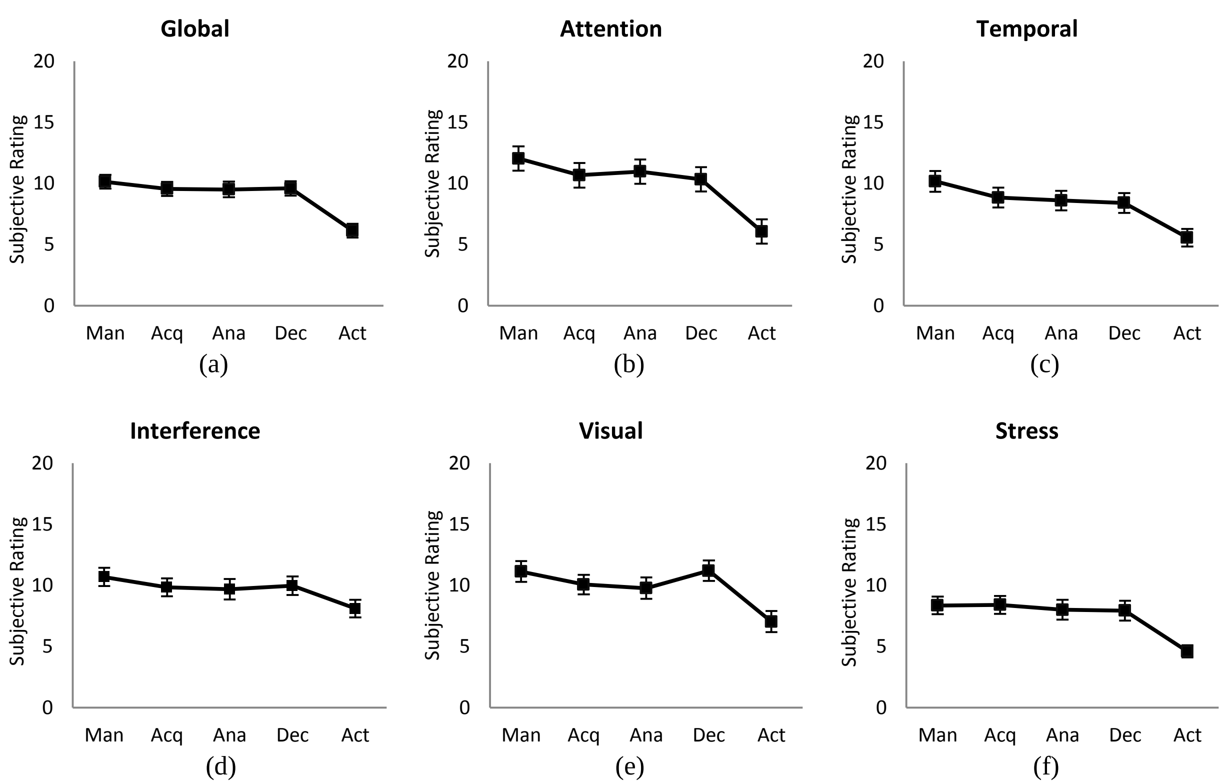


Figure 7-2. Means of DALI parameters for five ICMTA levels: Manual (Man), Information Acquisition (Acq), Information Analysis (Ana), Decision and action Selection (Dec), and Action Implementation (Act).

mode (M = 6.1, SE = 0.6), as indicated by pairwise analysis (Table 7-1). Means of the remaining ICMTA modes were between M= 9.5 (SE = 0.6) and M = 10.1 (SE = 0.6), showing no significant differences with one another. Analysis of the details of the results for each factor would allow better understanding of the components of the global workload score.

Table 7-1. Multiple comparison table on global mental workload of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.6	0.4	102	1.54	0.126	-1.48	0.18
	Ana	0.7	0.4	102	1.61	0.110	-1.52	0.16
	Dec	0.6	0.4	102	1.46	0.148	-1.44	0.22
	Act	4.1	0.4	102	9.55	<.001	-4.99	-3.27
Acq	Ana	0.0	0.4	102	0.08	0.938	-0.79	0.86
	Dec	0.0	0.4	102	0.09	0.929	-0.85	0.78
	Act	3.5	0.4	102	8.18	<.001	2.64	4.33
Ana	Dec	0.1	0.4	102	0.17	0.868	-0.89	0.75
	Act	3.5	0.4	102	8.07	<.001	-4.30	-2.60
Dec	Act	3.5	0.4	102	8.27	<.001	-4.37	-2.68

7.2.2 Attentional demand

There was a main effect of ICMTA on *attentional demand*, F (4, 108) = 21.39, p <.001, ω^2_p = .352, but no significant effects of VSTA or VSTA × ICMTA automation interaction were observed, p > .05. Figure 7-2b demonstrates results of attentional demand in different ICMTA levels. According to the pairwise analysis (Table 7-2) Action Implementation mode (M = 6.1, SE = 0.9) of ICMTA showed a different effect compared to the other ICMTA levels. In Action Implementation mode, the attention required by the task was less than the attention required during conditions when the operator was involved in the task loop. Attentional demand required by the Manual (M = 12.0, SE = 0.7) and Decision and Action Selection (M = 10.3, SE = 0.8) modes also were significantly different. Participants of the study gave a higher score for Manual mode, indicating it as the most demanding situation.

Table 7-2. Multiple comparison table on Attentional demand of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	1.4	0.7	108	1.95	0.054	-2.76	0.03
	Ana	1.1	0.7	108	1.52	0.132	-2.46	0.33
	Dec	1.7	0.7	108	2.42	0.017	-3.09	-0.31
	Act	6.0	0.7	108	8.5	<.001	-7.36	-4.57
Acq	Ana	0.3	0.7	108	0.43	0.670	-1.69	1.09
	Dec	0.3	0.7	108	0.47	0.636	-1.06	1.73
	Act	4.6	0.7	108	6.55	<.001	3.21	5.99
Ana	Dec	0.6	0.7	108	0.9	0.369	-0.76	2.03
	Act	4.9	0.7	108	6.98	<.001	-6.29	-3.51
Dec	Act	4.3	0.7	108	6.08	<.001	-5.66	-2.87

7.2.3 Temporal demand

ICMTA level had a significant effect on *temporal demand*, F (4, 108) = 10.14, p < .001, ω^2_p = .196, with *temporal demand* generally decreasing as the ICMTA level increased (Figure 7-2c). Based on the post hoc analysis, as shown in table 7-3, Information Acquisition (M = 8.8, SE = 0.8), Information Analysis (M = 8.6, SE = 0.8), and Decision and Action Selection (M = 8.4, SE = 0.8) modes had similar effect on the *temporal demand* of operators. Action Implementation mode enacted the lowest *temporal demand* with the mean of M = 5.6 (SE = 0.7). By contrast, the Manual mode (M =10.2, SE = 0.9) caused the highest *temporal demand*. The main effect of VSTA was not significant, nor was the VSTA × ICMTA interaction effect on *temporal demand*, p > .05.

Table 7-3. Multiple comparison table on temporal demand of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	1.3	0.7	108	1.78	0.077	-2.82	0.15
	Ana	1.6	0.7	108	2.10	0.039	-3.05	-0.08
	Dec	1.8	0.7	108	2.36	0.020	-3.25	-0.28
	Act	4.6	0.7	108	6.15	<.001	-6.08	-3.12
Acq	Ana	0.2	0.7	108	0.31	0.756	-1.25	1.72
	Dec	0.4	0.7	108	0.58	0.563	-1.05	1.92
	Act	3.3	0.7	108	4.37	<.001	1.78	4.75
Ana	Dec	0.2	0.7	108	0.27	0.790	-1.28	1.68
	Act	3.0	0.7	108	4.06	<.001	-4.52	-1.55
Dec	Act	2.8	0.7	108	3.79	0.000	-4.32	-1.35

7.2.4 Interference

Disturbance of the driver's state and consequences on the driving activity when conducting the driving activity simultaneously with the control and monitoring of the implement is reflected in the *interference* component of the DALI questionnaire. The effect of ICMTA on *interference* was significant, F (4, 104) = 4.49, p < .001, ω^2_p = .097. Results of the study showed similar levels of interference for the first four levels of ICMTA (Figure 7-2d). Results of pairwise analysis are shown in Table 7-4. Manual (M =10.7, SE = 0.8), Information Acquisition (M = 9.8, SE = 0.7), Information Analysis (M = 9.7, SE = 0.8), and Decision and Action selection (M = 10.0, SE = 0.8) modes resulted in moderate levels of interference with the driving task. Action Implementation mode (M = 8.1, SE = 0.7) in this case, had a significantly different effect by imposing the least interference on the driving task. ANOVA indicated no main effect of VSTA or VSTA × ICMTA interaction, p > .05.

Table 7-4. Multiple comparison table on interference of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.7	0.6	104	1.20	0.234	-1.93	0.48
	Ana	0.8	0.6	104	1.33	0.186	-2.06	0.41
	Dec	0.6	0.6	104	0.98	0.330	-1.80	0.61
	Act	2.5	0.6	104	4.01	0.000	-3.68	-1.25
Acq	Ana	0.1	0.6	104	0.16	0.872	-1.12	1.32
	Dec	0.1	0.6	104	0.22	0.825	-1.32	1.06
	Act	1.7	0.6	104	2.86	0.005	0.53	2.94
Ana	Dec	0.2	0.6	104	0.38	0.706	-1.45	0.99
	Act	1.6	0.6	104	2.65	0.009	-2.86	-0.41
Dec	Act	1.9	0.6	104	3.08	0.003	-3.07	-0.67

7.2.5 Visual demand

ICMTA levels caused significant main effect on *visual demand*, F (4, 108) = 8.51, p < .001, ω^2_p = .167. Figure 7-2e illustrates means of visual demand in different ICMTA levels. Similar to *attentional demand,* Action Implementation mode (M = 7.0, SE = 0.9) showed significantly different effect, being the least demanding condition (Table 7-5). The rest of the ICMTA levels resulted in moderate to above-moderate values (Manual (M =11.1, SE = 0.8), Information Acquisition (M = 10.1, SE = 0.8), Information Analysis (M = 9.8, SE = 0.9), and Decision and Action selection (M = 11.2, SE = 0.9)), but they were not significantly different from one another. The main effect of VSTA was not significant, nor was the VSTA × ICMTA interaction, $p > .05$.

Table 7-5. Multiple comparison table on visual demand of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	1.1	0.8	108	1.30	0.196	-2.69	0.56
	Ana	1.4	0.8	108	1.67	0.099	-2.99	0.26
	Dec	0.1	0.8	108	0.08	0.935	-1.56	1.69
	Act	4.1	0.8	108	5.00	<.001	-5.73	-2.47
Acq	Ana	0.3	0.8	108	0.37	0.715	-1.33	1.93
	Dec	1.1	0.8	108	1.38	0.170	-2.76	0.49
	Act	3.0	0.8	108	3.70	0.000	1.41	4.66
Ana	Dec	1.4	0.8	108	1.75	0.083	-3.06	0.19
	Act	2.7	0.8	108	3.33	0.001	-4.36	-1.11
Dec	Act	4.2	0.8	108	5.08	<.001	-5.79	-2.54

7.2.6 Situational stress

ANOVA showed a significant main effect of ICMTA on *situational stress*, F (4, 99) = 21.41, $p < .001$, $\omega^2_p = .352$ .Considering the *situational stress* caused by ICMTA levels, all of the automation types resulted in below-moderate levels of *stress* on operators (Figure 7-2f). In this case, the Action Implementation mode (M = 4.6, SE = 0.5) was the least stressful situation for the subjects of the study as indicated by the multiple comparison table (Table 7-6). The first four ICMTA levels did not show significant differences with one another (Manual (M =8.4, SE = 0.7), Information Acquisition (M = 8.4, SE = 0.7), Information Analysis (M = 8.0, SE = 0.8), and Decision and Action selection (M = 7.9, SE = 0.8). No main effect of VSTA or VSTA × ICMTA interaction was observed, $p > .05$.

Table 7-6. Multiple comparison table on situational stress of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.1	0.5	99	0.22	0.829	-1.10	0.88
	Ana	0.5	0.5	99	0.95	0.343	-1.49	0.52
	Dec	0.6	0.5	99	1.11	0.269	-1.57	0.44
	Act	4.0	0.5	99	7.81	<.0001	-5.06	-3.01
Acq	Ana	0.4	0.5	99	0.75	0.453	-0.61	1.36
	Dec	0.5	0.5	99	0.91	0.363	-0.53	1.44
	Act	3.9	0.5	99	7.81	<.0001	2.93	4.92
Ana	Dec	0.1	0.5	99	0.16	0.876	-0.92	1.08
	Act	3.5	0.5	99	6.88	<.0001	-4.57	-2.53
Dec	Act	3.5	0.5	99	6.74	<.0001	-4.49	-2.45

7.3 Performance

7.3.1 Reaction time

In the main experiment, it was found that the manual steering task increased the average reaction time, similar to the findings from the preliminary experiment. The average reaction time in auto-steer mode was 0.67 s (SE = 0.12), while in manual steering mode it was 1.01 s (SE = 0.15). By eliminating results from the Action Implementation mode, in which operators were not involved in the task loop, the average reaction times were 1.96 s (SE = 0.11) and 1.71 s (SE = 0.12) for manual and automatic steering modes respectively. Figure 7-3 shows means of reaction time for different VSTA and ICMTA levels.

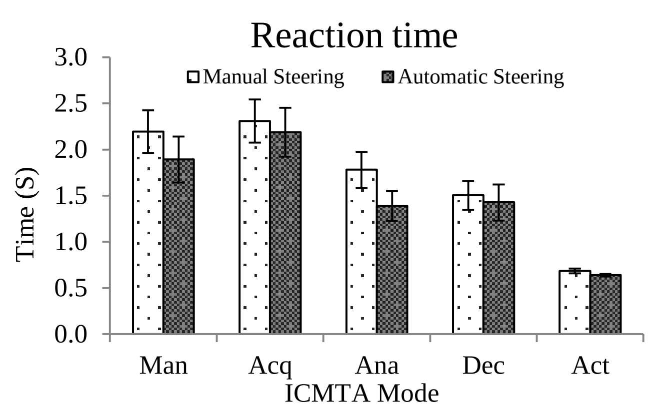


Figure 7-3. Reaction time of operators in different ICMTA and VSTA modes.

The ANOVA did not show a main effect of VSTA or interaction of VSTA × ICMTA on reaction time, $p > .05$, despite the lower values for reaction time in automatic steering mode in all of ICMTA levels. There was a significant main effect of ICMTA support on reaction time, F (4, 94) = 68.79, $p < .001$, $\omega^2_p = .670$. In general, results showed a decreasing trend for reaction time as the level of ICMTA increased. The highest reaction time was observed in Information Acquisition mode (M = 2.25, SE = 0.17 s), however, the table of differences of least square means (Table 7-7) did not show any significant differences between Manual (M = 2.04, SE = 0.17 s) and Information Acquisition modes. Information Analysis (M =1.58, SE = 0.13 s), and Decision and Action Selection (M = 1.46, SE = 0.12 s) modes had similar effects on reaction time, but with values significantly lower than Manual and Information Acquisition modes. Action Implementation mode resulted in the lowest reaction time (M = 0.66, SE = 0.02 s), presenting significant differences with other ICMTA modes.

Table 7-7. Multiple comparison table on reaction time of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.12	0.07	94	1.76	0.082	-0.02	0.26
	Ana	0.22	0.07	94	3.27	0.002	-0.36	-0.09
	Dec	0.31	0.07	94	4.55	<.001	-0.45	-0.18
	Act	0.96	0.07	94	13.79	<.001	-1.10	-0.82
Acq	Ana	0.35	0.07	94	4.81	<.001	0.20	0.49
	Dec	0.43	0.07	94	6.07	<.001	0.29	0.58
	Act	1.08	0.07	94	14.95	<.001	0.94	1.23
Ana	Dec	0.09	0.07	94	1.24	0.217	-0.05	0.23
	Act	0.74	0.07	94	10.25	<.001	-0.88	-0.59
Dec	Act	0.65	0.07	94	8.99	<.001	-0.79	-0.51

7.3.2 Number of errors

Means of the number of errors that participants made in different VSTA and ICMTA conditions are shown in Figure 7-4. In total, operators made 126 errors (62 errors were related to WT, 58 for WD and 6 for WP). Manual steering resulted in 76 failures and automatic steering resulted in only 50 failures, however, the difference was not significant, so there was no main effect of VSTA on the number of errors, $p > .05$. ICMTA support showed a significant effect on the number of failures, F (4, 108) = 13.55, $p < .001$, $\omega^2_p = .251$). Likewise, the ICMTA × VSTA interaction effect was significant, F (4, 108) = 3.04, $p = .01$, $\omega^2_p = .052$).

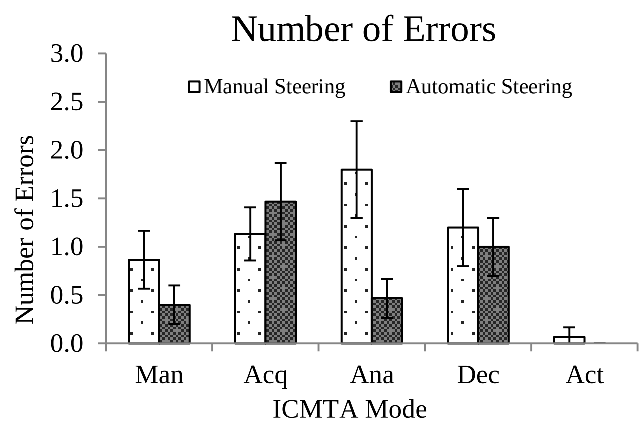


Figure 7-4. Number of errors made by operators in different ICMTA and VSTA modes.

Post hoc analysis (Table 7-8) indicated that Action Implementation mode (M = 0.03, SE = 0.03) significantly reduced the number of errors. A low number of errors were also observed in the Manual condition (M = 0.63, SE = 0.19). Both of these conditions showed significant differences with each other and the rest of the ICMTA

Table 7-8. Multiple comparison table on number of errors of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.36	0.10	108	3.49	0.001	0.16	0.56
	Ana	0.19	0.10	108	1.82	0.071	-0.02	0.39
	Dec	0.25	0.10	108	2.42	0.017	0.05	0.45
	Act	0.33	0.10	108	3.15	0.002	-0.53	-0.12
Acq	Ana	0.17	0.10	108	1.66	0.100	-0.03	0.38
	Dec	0.11	0.10	108	1.06	0.290	-0.09	0.31
	Act	0.68	0.10	108	6.64	<.001	0.48	0.89
Ana	Dec	0.06	0.10	108	0.6	0.552	-0.27	0.14
	Act	0.51	0.10	108	4.98	<.001	-0.72	-0.31
Dec	Act	0.58	0.10	108	5.58	<.001	-0.78	-0.37

modes, except for Information Analysis mode in case of Manual mode. Information Acquisition (M = 1.30, SE = 0.23), Information Analysis (M = 1.13, SE = 0.29), and Decision and Action Selection (M = 1.10, SE = 0.22) modes resulted in similar numbers of errors, showing no significant differences with one another. The ICMTA × VSTA interaction was due to the fact that subjects made fewer errors in Information Analysis Support mode while driving with the auto-steer system (M = 0.47, SE = 0.24) compared to manual steering (M = 1.8, SE = 0.49).

7.4 HRV

ANOVA of HRV parameters indicated no main effect of VSTA, $p > .05$. ICMTA effect was found on some of the time and frequency domain parameters. The VSTA × ICMTA interaction effect only was observed on two frequency domain parameters. In the following, only the HRV parameters with significant results are reported.

7.4.1 Minimum RR Interval

Minimum (min) RR Interval is the shortest interval between consecutive heartbeats during 10 min of driving. This parameter was affected by ICMTA, F (4, 99) = 4.09, $p < .01$, $\omega^2_p = .102$. The changes in min RR interval in different ICMTA levels are shown in Figure 7-5. Min RR interval decreased as the level of ICMTA increased, except for the Action Implementation mode (M = 636, SE = 12 ms) that resulted in longer min RR intervals than Decision and Action Selection mode (M = 630, SE = 14 ms). Pairwise analysis did not reveal any significant differences among Manual (M = 656, SE = 13 ms), Information Acquisition (M = 649, SE = 13ms), Information Analysis (M = 645, SE = 13 ms) and Decision and Action Selection modes (Table 7-9). Furthermore, no significant difference was found between Decision and Action Selection mode and Action

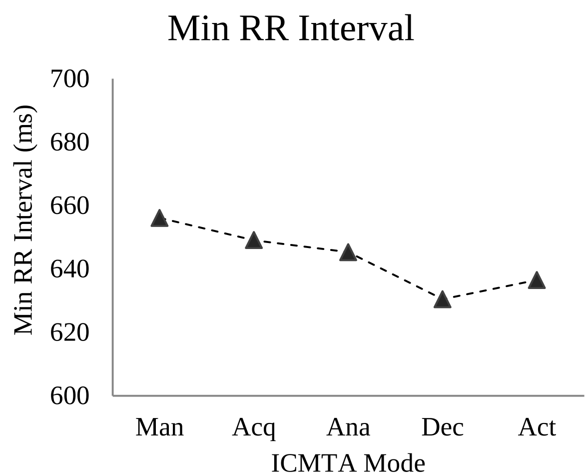


Figure 7-5. Min RR intervals for different ICMTA modes.

Implementation mode. Decision and Action selection mode showed differences with Manaual and Information Acquisition Modes. Another signifiant difference was observed between Manual and Action Implementation Modes. The main effect of VSTA was not significant, nor was the VSTA × ICMTA interaction effect on *temporal demand*, $p > .05$.

Table 7-9. Multiple comparison table on Min RR intervals of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	2E-06	2E-05	99	0.13	0.901	-0.00004	0.00003
	Ana	3E-05	2E-05	99	1.59	0.115	-0.00006	0.00001
	Dec	6E-05	2E-05	99	3.47	0.001	-0.00009	-0.00003
	Act	3E-05	2E-05	99	2.00	0.049	-0.00007	0.00000
Acq	Ana	3E-05	2E-05	99	1.45	0.152	-0.00001	0.00006
	Dec	6E-05	2E-05	99	3.3	0.001	0.00002	0.00009
	Act	3E-05	2E-05	99	1.86	0.065	0.00000	0.00001
Ana	Dec	3E-05	2E-05	99	1.85	0.068	0.00000	0.00007
	Act	7E-06	2E-05	99	0.41	0.682	-0.00004	0.00003
Dec	Act	3E-05	2E-05	99	1.44	0.153	-0.00001	0.00006

7.4.2 Max/min RR intervals ratio

The effect of ICMTA on Max/min RR intervals ratio was significant, F (4, 98) = 3.94, $p < .01$, $\omega^2_p = .119$. Max/min RR intervals ratio in different ICMTA levels is shown in Figure 7-6. The general trend was increasing by the level of ICMTA. No significant difference was found between Manual (M = 1.58, SE = 0.03) and Information

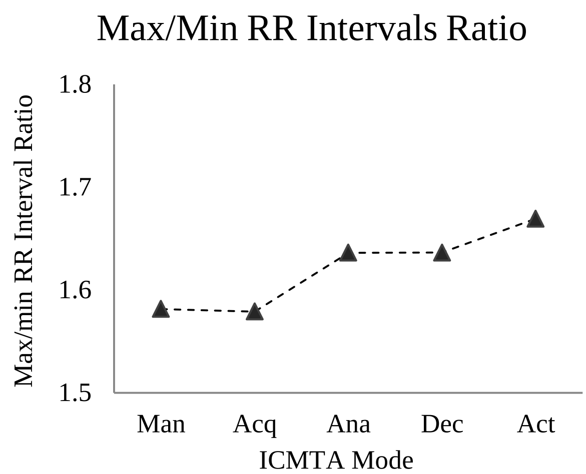


Figure 7-6. Max/min RR intervals ratio in different ICMTA modes.

Acquisition (M = 1.58, SE = 0.03) modes, according to the multiple comparisons table (Table 7-10). Similarly, Information Analysis (M = 1.64, SE = 0.04) and Decision and Action Selection (M = 1.64, SE = 0.03) modes had similar effects on the max/min ratio. Action Implementation mode (M = 1.67, SE = 0.04) resulted in the highest value. No main effect of VSTA or VSTA × ICMTA interaction was observed, $p > .05$.

Table 7-10. Multiple Comparison Table on Max/min RR intervals of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.00	0.01	98	0.21	0.831	-0.02	0.02
	Ana	0.02	0.01	98	2.09	0.039	-0.04	0.00
	Dec	0.02	0.01	98	1.99	0.049	-0.04	0.00
	Act	0.04	0.01	98	3.38	0.001	-0.06	-0.01
Acq	Ana	0.02	0.01	98	2.02	0.046	0.00	0.04
	Dec	0.02	0.01	98	2.11	0.037	0.00	0.04
	Act	0.03	0.01	98	3.17	0.002	0.01	0.05
Ana	Dec	0.00	0.01	98	0.23	0.822	-0.02	0.02
	Act	0.02	0.01	98	2.00	0.048	-0.03	0.01
Dec	Act	0.02	0.01	98	2.04	0.044	-0.04	0.01

7.4.3 LF

The changes in the 0.1 Hz component of HRV are shown in Figure 7-7. Manual steering mode resulted in lower values for LF compared to automatic steering mode, however, the differences were not statistically significant, $p > .05$. ICMTA showed a

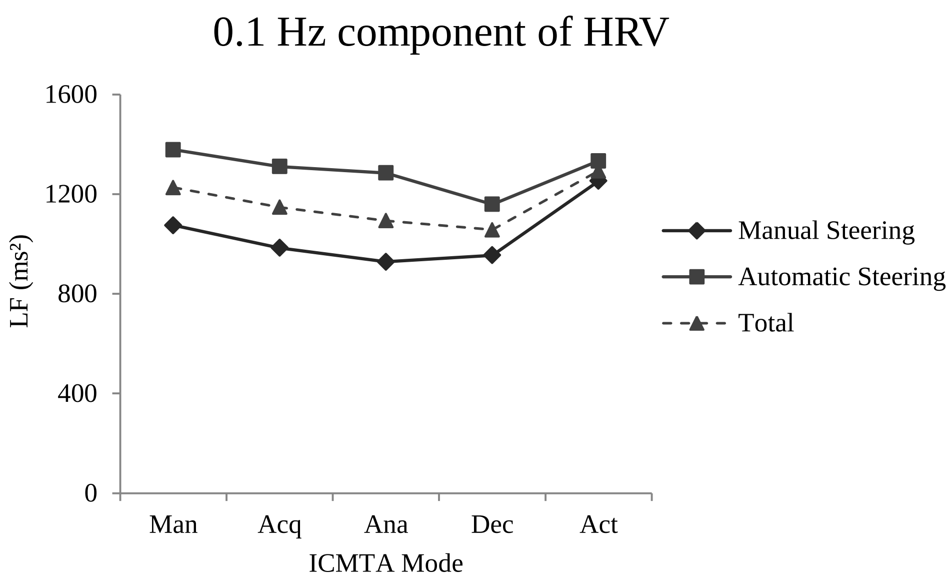


Figure 7-7. The 0.1 Hz component of HRV for different VSTA and ICMTA conditions.

significant effect on LF, F (4, 98) = 3.32, p = .014, ω^2_p = .079. It was observed that by increasing the ICMTA level, the 0.1 Hz component decreased until the Decision and Action Selection mode. The Action Implementation mode increased the LF value. Multiple comparison table (Table 7-11) did not reveal any differences among Manual (M = 1226, SE = 149 ms²), Information Acquisition (M = 1147, SE = 131 ms²), Information

Table 7.11. Multiple Comparison Table on the 0.1 Hz component of HRV of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.02	0.05	98	0.4	0.690	-0.08	0.13
	Ana	0.01	0.05	98	0.19	0.847	-0.12	0.10
	Dec	0.06	0.05	98	1.09	0.278	-0.16	0.05
	Act	0.13	0.05	98	2.38	0.019	0.02	0.23
Acq	Ana	0.03	0.05	98	0.59	0.553	-0.07	0.14
	Dec	0.08	0.05	98	1.49	0.139	-0.03	0.18
	Act	0.11	0.05	98	1.99	0.050	-0.21	0.00
Ana	Dec	0.05	0.05	98	0.9	0.372	-0.06	0.15
	Act	0.14	0.05	98	2.57	0.012	0.03	0.24
Dec	Act	0.19	0.05	98	3.48	0.001	0.08	0.29

Analysis (M = 1093, SE = 126 ms²), and Decision and Action Selection (M = 1056, SE =135 ms²) modes. On the other hand, Action Implementation mode (M = 1293, SE = 172 ms²) had a significantly different effect compared to the other ICMTA modes.

VSTA × ICMTA interaction was significant in case of the 0.1 Hz component of HRV, F (4, 98) = 2.63, p = .038, ω^2_p = .046. In manual steering mode, Action Implementation caused a significantly higher value compared to the rest of ICMTA modes. In auto-steer condition, only Decision and Action Selection mode showed significant differences with other ICMTA modes, resulting in the lowest value.

7.4.4 LF/HF ratio

Figure 7-8 illustrates variations in LF/HF ratio in different automation conditions. The LF/HF ratio remained at similar levels for all of the driving blocks in automatic steering mode. A dramatic change was observed in the manual steering mode. No main effect of VSTA was observed on LF/HF ratio, p > .05. The ANOVA revealed a significant main effect of ICMTA, F (4, 95) = 8.67, p < .001, ω^2_p = .077, and a VSTA × ICMTA

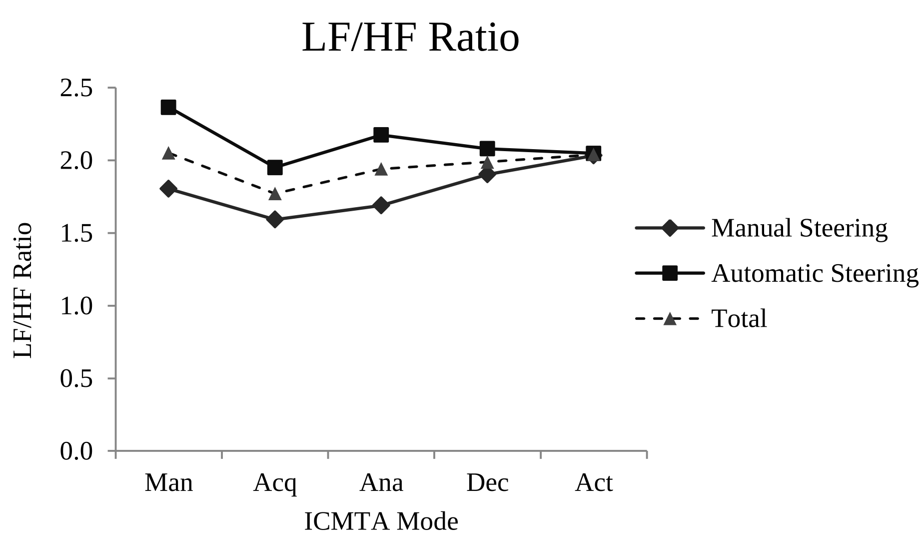


Figure 7-8. The LF/HF ratio for different VSTA and ICMTA conditions.

interaction effect, F (4, 95) = 5.28, p < .001, ω^2_p = .086, for LF/HF ratio. Table 7-12 demonstrates the results of post hoc analysis. The highest LF/HF ratio was observed in Action Implementation mode (M = 0.54, SE = 0.11) while the Information Acquisition mode (M = 0.41, SE = 0.12) resulted in the lowest ratio. Aside from the Manual mode (M = 0.51, SE = 0.14), the trend was increasing LF/HF ratio as the level of ICMTA increased. The effect of Information Acquisition mode was significantly different from all of the other automation modes. Action Implementation mode also showed significant differences with all of the ICMTA modes except for the Manual mode.

Table 7-12. Multiple comparison table on LF/HF ratio of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.2	0.0	95	4.04	0.000	-0.26	-0.09
	Ana	0.1	0.0	95	1.81	0.073	-0.17	0.01
	Dec	0.1	0.0	95	1.18	0.241	-0.14	0.03
	Act	0.1	0.0	95	1.34	0.184	-0.03	0.14
Acq	Ana	0.1	0.0	95	2.26	0.026	-0.18	-0.01
	Dec	0.1	0.0	95	3.01	0.003	-0.20	-0.04
	Act	0.2	0.0	95	5.57	<.001	-0.31	-0.15
Ana	Dec	0.0	0.0	95	0.7	0.487	-0.11	0.05
	Act	0.1	0.0	95	3.27	0.002	0.05	0.22
Dec	Act	0.1	0.0	95	2.64	0.010	0.03	0.19

7.4.5 I_{PNS}

Normalized power of HF, I_{PNS}, only showed changes to variations in level of ICMTA, F (4, 89) = 2.87, p = .03, ω^2_p = .058. Information Acquisition mode (M = 0.39, SE = 0.02) resulted in higher I_{PNS} compared to Manual mode (M = 0.37, SE = 0.03). By increasing the level of support after Information Acquisition mode, a decreasing trend could be observed (Figure 7-9). Despite the variations in I_{PNS}, post-hoc analysis (Table 7-13) only showed significant differences between Information Acquisition mode and the rest of the ICMTA levels, except for Decision and Action Selection mode (M = 0.37, SE = 0.03). In other words, no significant differences were found among Manual,

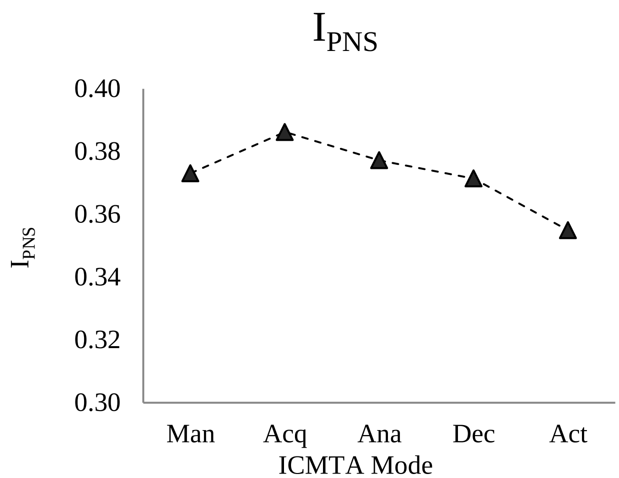


Figure 7-9. Index of Parasympathetic Nervous System (I_{PNS}) for different ICMTA modes.

Information Analysis (M = 0.38, SE = 0.03), Decision and Action Selection and Action Implementation (M = 0.35, Se = 0.04) modes. The main effect of VSTA and VSTA × ICMTA interaction were not significant, $p > .05$.

Table 7-13. Multiple Comparison Table on I_{PNS} of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.1	0.0	89	2.37	0.020	0.01	0.17
	Ana	0.0	0.0	89	0.11	0.916	-0.07	0.08
	Dec	0.0	0.0	89	0.69	0.495	-0.05	0.10
	Act	0.0	0.0	89	0.86	0.393	-0.11	0.04
Acq	Ana	0.1	0.0	89	2.29	0.025	0.01	0.16
	Dec	0.1	0.0	89	1.72	0.089	-0.01	0.14
	Act	0.1	0.0	89	3.19	0.002	0.05	0.20
Ana	Dec	0.0	0.0	89	0.57	0.567	-0.10	0.05
	Act	0.0	0.0	89	0.97	0.337	-0.12	0.04
Dec	Act	0.1	0.0	89	1.51	0.135	-0.14	0.02

7.5 Situation awareness

Results of situation awareness and its components in different VSTA modes from the main experiment are shown in Figure 7-10. SART-combined scores ranged from 6.8 to 31.8, with 31.8 denoting high situation awareness. Automatic steering mode resulted in slightly higher situation awareness mainly due to lower attentional demand.

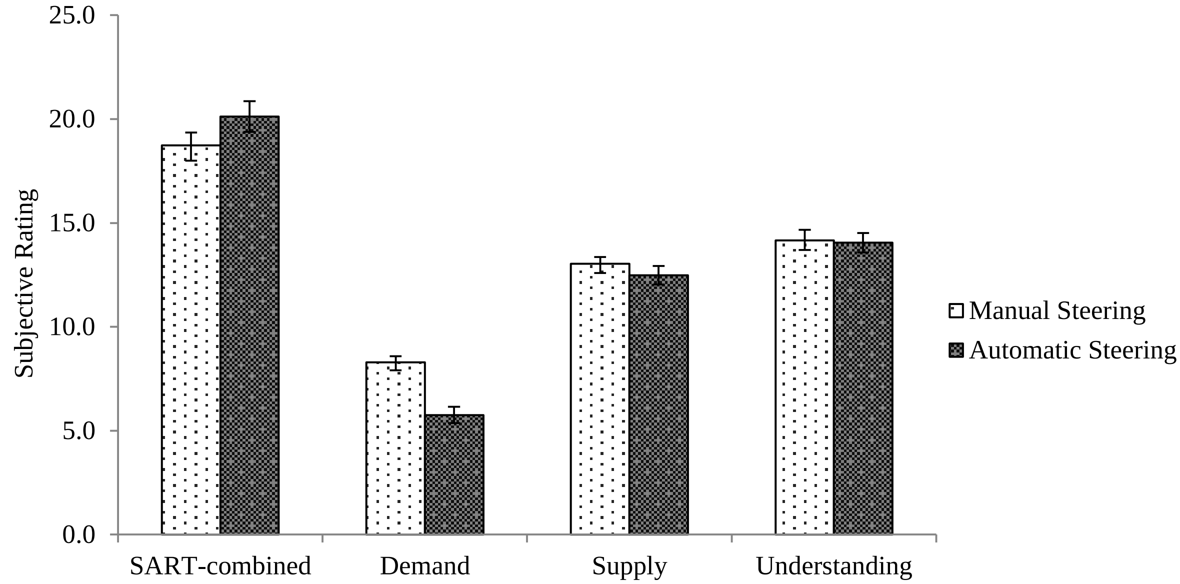


Figure 7-10. Overall Situation awareness and its components for different VSTA modes for the main experiment.

ANOVA revealed no VSTA or VSTA × ICMTA interaction effects on SART-combined, $p > .05$. There was a significant effect of ICMTA on SART-combined $F(4, 107) = 5.64$, $p < .01$, ω2p = 0.109. As shown in Figure 7-11, Situation awareness increased as automation was applied to higher levels of information processing functions, however, there was a sudden drop in the Action Implementation mode.

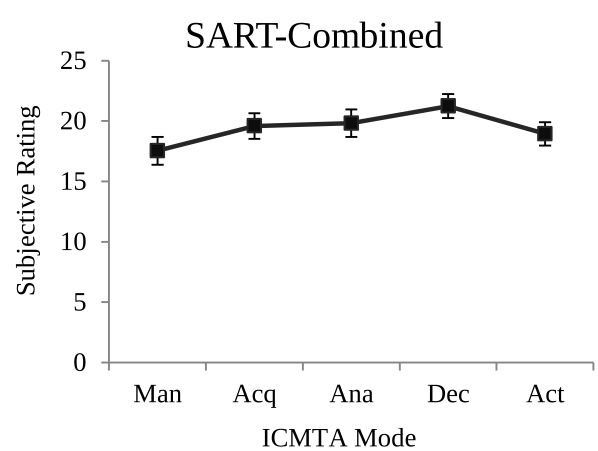


Figure 7-11. Situation awareness rating for different ICMTA modes

According to the post-hoc analysis (Table 7-14), a significant difference was found between Decision and Action Selection mode (M = 21.2, SE = 1.4) and all of the other ICMTA support modes except for the Information Analysis mode (M = 19.8, SE = 1.4). Furthermore, Manual (M = 17.5, SE = 1.4) and Information Analysis modes showed different effects from one another.

Table 7-14. Multiple comparison table on SART-combined of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	2.1	0.8	107	2.56	0.012	0.46	3.64
	Ana	2.3	0.8	107	2.85	0.005	0.70	3.88
	Dec	3.7	0.8	107	4.62	<.001	2.11	5.29
	Act	1.4	0.8	107	1.74	0.085	-0.20	3.02
Acq	Ana	0.2	0.8	107	0.29	0.769	-1.83	1.35
	Dec	1.7	0.8	107	2.06	0.042	-3.24	-0.06
	Act	0.6	0.8	107	0.79	0.432	-0.97	2.25
Ana	Dec	1.4	0.8	107	1.76	0.081	-3.00	0.18
	Act	0.9	0.8	107	1.08	0.282	-2.49	0.73
Dec	Act	2.3	0.8	107	2.82	0.006	-3.90	-0.68

7.5.1 Demand on attentional resources

Figure 7-12 shows demand subjective ratings for different VSTA and ICMTA levels, as well as average ratings for ICMTA levels. Overall, the TAS operating task demanded below moderate levels of attentional resources. The main effect of VSTA on this parameter was significant, F (1, 28) = 5.32, p = .03, ω^2_p = .216, as was the main effect of ICMTA, F (4, 108) = 16.91, p < .001, ω^2_p = .299. According to pairwise analysis (Table 7-15), automatic steering (M = 5.8, SE = 1.0) resulted in lower task demand compared to manual steering mode (M = 8.3, SE = 1.0). Considering the average ratings in Figure 7-12, it was found that by increasing the level of automation of the air seeder control and monitoring task, operators experienced lower task attentional demands. The Manual mode (M = 9.0, SE = 0.8) was the most demanding situation. It was observed

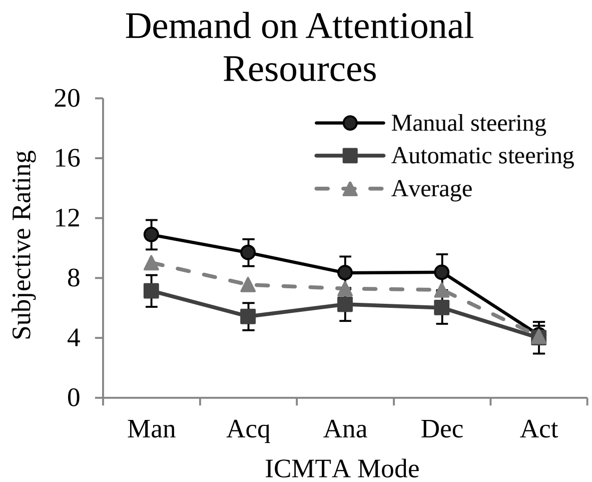


Figure 7-12. Demand subjective ratings for different VSTA and ICMTA levels.

that subjects gave higher scores to the three subscales of demand (instability, complexity and variability) in this mode. In fact, subjects thought that the ICMTA mode with lowest automation support level was highly complex, unstable and varying. Following a decreasing trend, however, the three ICMTA modes of Information Acquisition (M = 7.6, SE = 0.7), Information Analysis (M = 7.3, SE = 0.8), and Decision and Action Selection (M = 7.2, SE = 0.9) did not show any different impacts from one another. The Action Implementation mode (M = 4.1, SE = 0.6), as the highest level of ICMTA, imposed the lowest level of attentional demand.

Table 7-15. Multiple comparison table on demand on attentional resources of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	1.5	0.6	108	2.36	0.020	-2.68	-0.24
	Ana	1.7	0.6	108	2.79	0.006	-2.95	-0.50
	Dec	1.8	0.6	108	2.95	0.004	-3.04	-0.60
	Act	4.9	0.6	108	7.96	<.001	-6.14	-3.69
Acq	Ana	0.3	0.6	108	0.43	0.667	-0.96	1.49
	Dec	0.4	0.6	108	0.58	0.561	-0.86	1.58
	Act	3.5	0.6	108	5.6	<.001	2.23	4.68
Ana	Dec	0.1	0.6	108	0.15	0.880	-1.13	1.32
	Act	3.2	0.6	108	5.16	<.001	-4.41	-1.97
Dec	Act	3.1	0.6	108	5.01	<.001	-4.32	-1.87

There was a significant interaction between VSTA and ICMTA over *demand on attentional resources*, F (4, 108) = 3.34, p = .01, ω^2_p = .059. It was found that manual steering generally resulted in higher attentional demand for all of the ICMTA levels. Comparing the same levels of ICMTA at different VSTA levels, statistical analysis showed that only Manual modes and Information Acquisition modes had different impacts. In manual steering mode, a decreasing trend was observed in the subject's ratings as the level of ICMTA support increased. The lowest level ICMTA (i.e., the Manual mode) was the most demanding situation for operators (M = 10.9, SE = 1.0). Action Implementation mode, on the other hand, required the least level of attention (M = 4.2, SE = 0.6). Requiring similar amounts of attentional resources, subjects did not differentiate the three intermediate levels. Variations of attentional demand ratings in automatic steering mode followed a slightly different trend. Subjects reported similar demand for the first four levels of ICMTA. In this VSTA mode, Action Implementation (M = 4.0, SE = 1.1) remained as the least demanding situation.

7.5.2 Supply of attentional resources

In the case of *supply of attentional resources*, no VSTA effect was found, p > .05, but the ICMTA effect was significant, F (4, 102) = 28.43, p <.001, ω^2_p = .427. As shown in Figure 7-13, among all of the ICMTA levels, Action Implementation resulted in significantly lower ratings (M = 10.3, SE = 0.5). The average supply rating in Action Implementation mode was 21% less than the average rating for the rest of the ICMTA levels. Pairwise analysis (Table 7-16) did not demonstrate any differences among Manual (M = 13.4, SE = 0.5), Information Acquisition (M = 13.5, SE = 0.4), and Information Analysis (M = 13.1, SE= 0.5) modes. Decision and Action Selection mode (13.7 ± 0.5) showed differences with Information Analysis and Action Implementation modes. The VSTA × ICMTA interaction showed no significant effect on *supply of attentional resources*, $p > .05$.

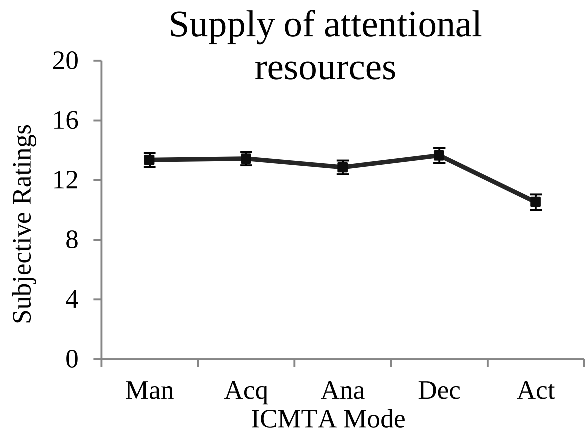


Figure 7-13. Subjective ratings of supply of attentional resources for different ICMTA levels.

Table 7-16. Multiple comparison table on supply of attentional resources of ICMTA levels.

Condition (I)	Condition (J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.1	0.3	102	0.27	0.786	-0.57	0.75
	Ana	0.5	0.3	102	1.45	0.151	-1.16	0.18
	Dec	0.3	0.3	102	0.92	0.360	-0.35	0.96
	Act	2.8	0.3	102	8.31	<.001	-3.49	-2.14
Acq	Ana	0.6	0.3	102	1.72	0.089	-0.09	1.25
	Dec	0.2	0.3	102	0.65	0.519	-0.87	0.44
	Act	2.9	0.3	102	8.57	<.001	2.24	3.58
Ana	Dec	0.8	0.3	102	2.36	0.020	-1.46	-0.13
	Act	2.3	0.3	102	6.78	<.001	-3.01	-1.65
Dec	Act	3.1	0.3	102	9.33	<.001	-3.79	-2.46

7.5.3 Understanding

Similar to the demand and supply, the ICMTA effect on *understanding* was significant, $F(4, 100) = 4.12$, $p = .004$, $\omega^2_p = .082$. As shown in Figure 7-14, participants reported better situational understanding as the ICMTA increased. This trend ended when the highest level of ICMTA (Action Implementation) was introduced. Surprisingly, participants ratings of *understanding* for the Mnual and Action Implementation modes were similar. As shown in Table 7-17, ratings of *understanding* in Action Implementation (M = 13.5, SE = 0.7) and Manual (M = 13.5, SE = 0.7) modes found to be significantly lower than for either Information Analysis (M = 14.6, SE = 0.6) or Decision and Action Selection (M = 14.7, SE = 0.5) support modes. No significant effect of VSTA or VSTA × ICMTA interaction was found on *understanding*, p > .05.

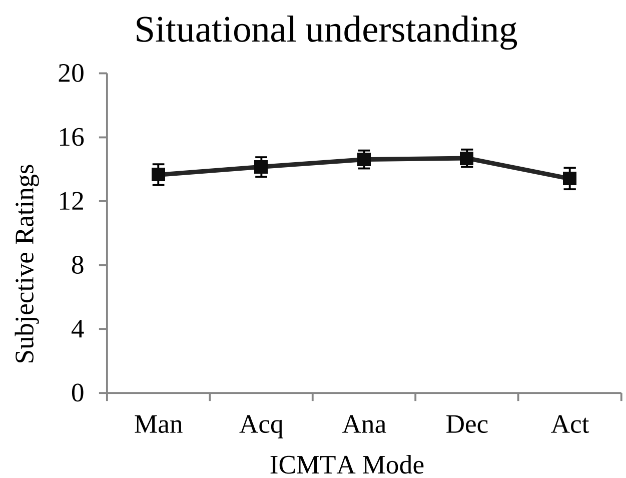


Figure 7-14. Subjective ratings for situational understanding for different ICMTA modes.

Table 7-17. Multiple comparison table on situational understanding of ICMTA levels.

Condition(I)	Condition(J)	Mean Difference |I-J|	Std. Error	df	|t|	Pr>t	95% Confidence Interval	
							Lower limit	Upper limit
Man	Acq	0.5	0.4	100	1.23	0.2202	-0.30	1.28
	Ana	1.0	0.4	100	2.41	0.0176	0.17	1.74
	Dec	1.0	0.4	100	2.58	0.0112	0.24	1.81
	Act	0.2	0.4	100	0.6	0.5514	-1.04	0.56
Acq	Ana	0.5	0.4	100	1.22	0.2265	-1.22	0.29
	Dec	0.5	0.4	100	1.39	0.1671	-1.30	0.23
	Act	0.7	0.4	100	1.88	0.0629	-0.04	1.50
Ana	Dec	0.1	0.4	100	0.19	0.8482	-0.83	0.68
	Act	1.2	0.4	100	3.09	0.0026	-1.96	-0.43
Dec	Act	1.3	0.4	100	3.24	0.0016	-2.04	-0.49

7.6 Correlation analysis

Correlation analyses were conducted using the Pearson coefficient in order to identify any significant relationships among subjective mental workload, subjective SA, performance and HRV. First, global scores of DALI, SART, reaction time, number of errors, and parameters of HRV were analysed. Next, components of DALI and SART were examined against performance and some HRV parameters. Using the same scale as for preliminary experiment (Table 6-10), the strength of positive or negative correlations were interpreted.

7.6.1 DALI, SART, Reaction time and Error

Table 7-18 summarises the results of correlation analysis of perceived workload, situation awareness, reaction time and number of errors from the main experiment. No perfect correlations were found among the results. Mental workload and situation awareness were negatively correlated. Mental workload showed positive correlations with reaction time and number of errors, while situation awareness did not have any correlations with these parameters. No correlations were found between HRV parameters and subjective scores of workload and situation awareness. The analysis indicated that reaction time and number of errors were positively correlated. Among all of the parameters, only reaction time showed correlations with HRV parameters. It was negatively correlated with LF and positively correlated with Max/Min ratio.

Table 7-18. Pearson correlations for perceived workload, situation awareness, reaction time and number of errors.

	Workload	**Situation Awareness**	**Reaction Time**	**Number of Errors**	**LF**
Situation Awareness	-0.26**				
Reaction Time	0.46**	-0.12			
Number of Errors	0.30**	0.02	0.44**		
LF	-0.13	-0.04	-0.21**	-0.08	
Max/Min Ratio	0.15	-0.03	0.21**	-0.02	-0.49**

** $p < 0.01$ level.
* $p < 0.05$ level.

7.6.2 DALI and Performance

The correlations between mental workload and performance parameters are shown in Table 7-19. There were positive correlations among all of the parameters. The only non-significant correlation was found between attentional demand and number of errors. Notable strong correlations were observed between attentional demand and visual demand, and temporal demand and stress. The strong correlations between the global workload score and its parameters occurred because the global workload score was derived from its parameters. In the case of reaction time, moderate correlations were found with attentional demand, stress and interference, and a strong correlation was found with temporal demand. Number of errors had a moderate correlation with situational stress and temporal demand, showing the effects of timing demand and level of stress while conducting the activity on performance of the subjects.

Table 7-19. Pearson correlations for workload components, reaction time and number of errors.

	Attention	**Visual**	**Stress**	**Temporal**	**Interference**
Visual	0.73**				
Stress	0.43**	0.37**			
Temporal	0.48**	0.44**	0.76**		
Interference	0.40**	0.29**	0.56**	0.59**	
Global	0.79**	0.74**	0.79**	0.82**	0.70**
Reaction Time	0.34**	0.18*	0.45**	0.51**	0.39**
Number of Errors	0.14	0.19*	0.32**	0.32**	0.27**

** $p < 0.01$ level.
* $p < 0.05$ level.

7.6.3 DALI and HRV

Correlation coefficients of workload and HRV parameters are shown in Table 7-20. Min RR interval only showed small correlations with attentional and visual demands. Max/Min RR intervals ratio and LF only showed small associations with interference component of DALI. LF/HF Ratio was correlated with all of the DALI components. All of these were considered small asociations except for the medium negative correlation with stress.

Table 7-20. Pearson correlations for workload and HRV components.

	Attention	Visual	Stress	Temporal	Interference
Min RR Interval	0.23*	0.25**	0.05	0.02	0.06
Max/Min Interval Ratio	-0.06	-0.16	0.10	0.12	0.23**
LF	-0.02	0.10	-0.03	-0.11	-0.23**
LF/HF Ratio	-0.29**	-0.19*	-0.34**	-0.18*	-0.25**
I_{PNS}	0.21*	0.16	0.29**	0.16	0.19*

** $p < 0.01$ level.
* $p < 0.05$ level.

7.6.4 DALI and SART

The correlation analysis of parameters of workload and situation awareness is presented in Table 7-21. The demand component of situation awareness showed significant positive associations with workload parameters. In situations where demand on attentional resources was higher, subjects reported higher mental workload. Supply of attentional resources also had significant positive correlations with attention and visual demand. The negative correlation between visual demand and the understanding component of situation awareness indicates that visual demand increases have negative effect on situational understanding. Understanding also showed positive correlations with stress and temporal demand. It seems that increasing stress and timing pressure caused better understanding.

Table 7-21. Pearson correlations of workload and situation awareness components.

	Attention	Visual	Stress	Temporal	Interference	Demand	Supply
Demand	0.43**	0.34**	0.47**	0.57**	0.39**		
Supply	0.43**	0.38**	-0.05	0.05	-0.07	0.21*	
Understanding	-0.13	-0.24**	0.34**	0.26**	0.19*	0.03	-0.02

** $p < 0.01$ level.
* $p < 0.05$ level.

7.6.5 SART, performance and HRV

In addition to the analyses mentioned above, correlations among components of SART, performance and HRV were calculated (Table 7-22). Between situation awareness and performance parameters, only one significant correlation existed between demand on attentional resources and reaction time. Min RR interval showed a minor positive

Table 7-22. Pearson correlations of parameters of SA, performance and HRV.

	Demand	Supply	Understanding	Reaction Time	Number of Errors
Reaction Time	0.33**	0.12	0.07		
Number of Errors	0.14	0.11	0.07	0.48**	
Min RR Interval	0.02	0.22*	-0.19*	0.07	0.05
Max/Min Interval Ratio	0.16	-0.03	0.26**	0.27**	0.06
LF	-0.17	0.21*	-0.22**	-0.05	0.03
LFHF	0.00	-0.07	-0.13	-0.12	-0.06
I_{PNS}	-0.01	0.11	0.14	0.14	0.05

** $p < 0.01$ level.
* $p < 0.05$ level.

correlation with supply of attentional resources and a minor negative correlation with understanding. Max/Min RR interval ratio also showed small positive correlations with understanding and reaction time. Similar to MinRR, LF showed small correlations with supply and understanding. The rest of the parameters did not demonstrate any significant correlations.

7.7 Post-experiment Questionnaire

For the main experiment, relatively more feedback were collected from participants compared to the pilot study. All of the participants accurately answered the first query. They reported that they spent, on average, 29% of their time looking at the mapping system, 43% at the air seeder information display, 19% at the visual scenery, 6% at the implement (i.e., monitors behind the cab), and 2% at other items. Drivers in manual steering condition spent 48% of their time on looking at the mapping system. As shown in Figure 7-15, this was much lower (10%) for drivers in automatic steering condition. Instead, participants in auto-steer mode allocated more time on monitoring the implement parameters on display (53% vs. 33%), gazing forward on field scenery (27% vs 11%) and looking at everything else (3% vs 1%). The time spent on looking back to monitor the implement units (i.e. monitors behind the cab) was slightly higher for participants in manual steering mode (7% vs. 5%).

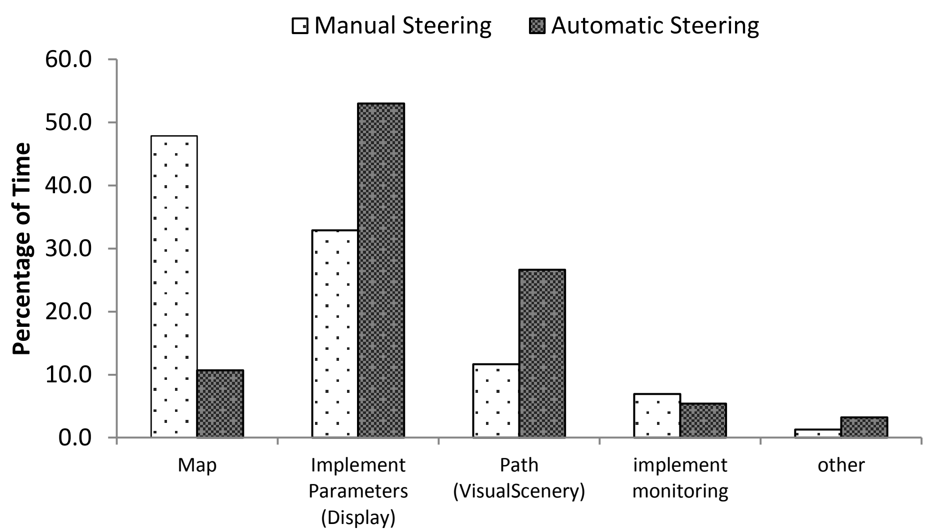


Figure 7-15. The average time spent on supervising various item in the simulator in different VSTA modes.

For the second question, most of the participants found it somewhat easy to locate items on the instrument console (i.e., 27% of participants found items easily, 63% of them found the task rather easy, and 10% of them found the task neither easy nor difficult). Figure 7-16 shows the ease of air seeder parameter finding for participants in different VSTA modes. As it can be seen, none of participants rated it difficult or very difficult. This result is due to the experience of participants with agricultural vehicles.

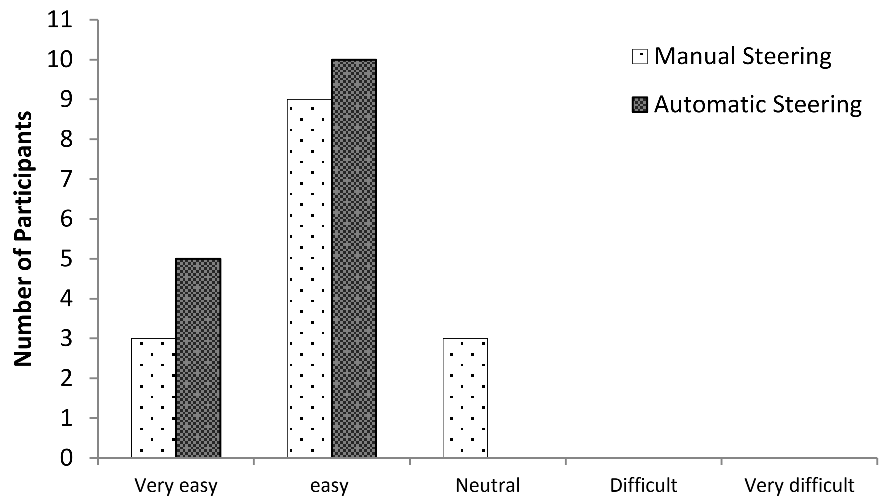


Figure 7-16. Ease of air seeder parameter finding on the console for participants.

In case of the third enquiry, most of the participants had a positive reaction to answering identical DALI and SART questionnaires after each driving period. Some of the participants found it interesting that they answered differently after different driving conditions. Some of the comments are noted below:

"good, because my perspective changed with the different levels."

"Good, knew what to think about and monitor personally."

"I feel that it was useful as it got me to really think about what changed between tests and how it changed my state of mind."

Just few participants answered neutral or found it repetitive. The following comments from participants are in regard to these views:

"I think my answers were similar for all of the questionnaires."

"A little repetitive"

"Neutral"

"Thought there was overlap in questions."

For the last question, a variety of responses were collected from participants. A few of them acknowledged the similarity of the air seeder display configuration with real displays. Comments included:

"I thought the simulator parameters for the air seeder was very easy to understand probably because I have worked with GPS air seeder equipment before."

"Very realistic simulator"

"Very well-done program and set up"

"Display was realistic and effective."

Four participants indicated that warning messages were very useful. Two comments are noted below:

"When the warning messages would say for example "decrease fan speed" it was helpful, as I did not have to check what the problem was, then find the control and then fix it. Much less intensive when the warning have detail."

"The error message helped gave time to find the parameters before they got to too low/high of a level."

Seven participants reported issues they confronted while driving the simulator regarding system response delay. Some of the comments are noted below:

"Visual scenery did not match with what you did while driving."

"Navigation map was frustrating at times with delays from steering."

The rest of the answers included suggestions for improving the simulator. All of the comments from participants to open-ended queries in the full scale experiment can be found in Appendix C.

Chapter 8

DISCUSSION

8.1 Introduction

The previous chapter reported the results of the the full scale experiment. This chapter reflects on the main findings of the research; Interpretations of the significant results, consistency of the results with the stated hypotheses, and correspondence of the results with the literature are provided. For the convenience of readers, each section in this chapter refers back to relevant tables or graphs from the chapter 7.

8.2 Mental workload

It was expected that mental workload would decrease as the level of automation support increased. This expectation was confirmed by subjective mental workload assessment in case of ICMTA. Generally, a decreasing trend was observed in global scores of mental workload as the level of automation increased (see Figure 7.2). This result is consistent with findings reported in previous studies of driver's mental workload (Gabaude et al. 2012; Pauzié et al. 2007; Petzoldt et al. 2011) that difficulty of an additional task causes higher subjective workload scores.

Although the VSTA effect was not significant, the ICMTA effect had a significant impact on mental workload scores. In the case of ICMTA, DALI was able to indicate two workload levels: i) high workload when subjects were involved in the task loop and ii) low workload during high automation conditions. The fact that the first four levels of ICMTA resulted in similar workload scores can be due to the following reasons. First, the physical workload might have a great impact on DALI scores in the experiment. From sensing (Manual and Information Acquisition modes) to deciding (Decision and Action

Selection mode), the physical workload remained equal as parameter adjustments were performed by operators. Next, automation support levels that were defined for these conditions practically do not make substantial differences with one another. Furthermore, it might be possible that DALI was not sensitive enough to identify small differences between these conditions.

With respect to the separate DALI components, only ICMTA resulted in significant differences in subjective rating. Similar to global mental workload scores, for all of the dimensions (i.e., attentional demand, temporal demand, interference, visual demand, and situational stress), higher scores were obtained when the drivers were involved in the task loop. The correlation analysis showed that all of these dimensions almost equally contributed to the global workload score. This result indicates the importance of including all of these dimensions in similar experimental conditions.

8.3 Performance

In the case of reaction time in the main experiment, the hypothesis was that the driver's performance would vary with level of automation. The hypothesis was confirmed in the case of both performance parameters (i.e., reaction time and number of errors). The VSTA effect was not significant, but in the case of ICMTA, increasing automation level was associated with lower reaction time and number of errors. This result confirms findings of previous studies (Gempton et al. 2013; Sethumadhavan 2009) by showing benefits to the performance of drivers due to automation.

The lowest reaction time and number of errors were achieved with the highest level of ICMTA support. In this mode, the automated system was responsible for parameter adjustment. It was expected that the average reaction time would be zero or very close to zero in this mode, but the response rate of the simulator computer and I/O boards did not allow for such a fast reaction. Furthermore, the Manual mode caused lower number of errors compared to Information Acquisition, Information Analysis, and Decision and Action Selection modes. This could be attributed to the individual's performance when driving the simulator. Some of the subjects made many errors while some of them did not make any errors.

Positive correlations among *global workload score*, reaction time and number of errors suggest benefits for automation in agricultural vehicles considering reduction in mental workload by automation level increase. Hwang et al. (2008) observed that subjects gave higher scores to a mental workload questionnaire when they made more mistakes.

8.4 HRV

It was expected that variations in the task automation would affect the HRV of drivers. Support for this hypothesis was found in spite of some contradictory results that could be seen in some cases. Five out of 15 HR and HRV parameters presented changes to VSTA and ICMTA variations.

As the level of ICMTA increased, Min RR interval decreased. Considering the reverse relationship between subjective mental workload and level of automation, the result from min RR interval is contradictory. Considering that max RR interval did not show changes to automation alterations, the max/min RR interval ratio can be a better parameter for assessing mental workload. A lower max/min RR interval ratio has shown lower parasympathetic activity (Nayem et al. 2013). Assuming the dynamically changing situation in the experiments, lower PNS activity indicates higher SNS activity and, therefore, higher mental workload.

The max/min RR interval ratio indicated three levels of workload: i) higher workload in Manual and Information Acquisition modes, ii) medium workload in Information Analysis and Decision Selection modes, and iii) lower workload in Action Implementation mode. There were distinct differences in the design of the information display for each of these three levels. For the first level (i.e., Manual and Information Acquisition modes), no messages were provided in the implement information display. The types of messages in the Information Analysis and Decision-making modes were alarming, making operators aware of projected errors. In Action Implementation automation, only informative messages were shown to operators, indicating adjustments that were made by the machine. Therefore, providing messages to operator and the type of message can have a great impact on mental workload. This result was partially in accordance with the subjective mental workload measurement in which only two levels of mental workload could be identified.

It was observed that the highest level of ICMTA (i.e., Action Implementation), which took the drivers out of the task loop, resulted in a higher LF value, indicating the lowest mental workload. In conditions where subjects were involved in the task-loop (i.e., Manual, Information Acquisition, Information Analysis, and Decision support modes), a decreasing trend (as opposed to the expected increasing trend) was observed in the 0.1 Hz component of HRV as the level of ICMTA increased, but the differences between the ICMTA modes were not statistically significant. Thus, the 0.1 Hz component of HRV could only identify two levels of mental workload in case of ICMTA, consistent with the results from the subjective mental workload measure in this study. This finding is in accordance with the studies that reported decreasing trend for the 0.1 Hz component of HRV by increased mental load (Di Marco et al. 2010; Mehler et al. 2011; Ramon et al. 2008).

LF/HF ratio and I_{PNS} showed inconsistencies. LF/HF ratio remained at a similar level for all of the ICMTA modes, except for a lower value in Information Acquisition mode. I_{PNS} values were also similar for all of ICMTA conditions, except for the higher value for the Information Acquisition mode.

The insensitivity of 0.1 Hz component of HRV, which is widely used for mental workload assessment, to mental workload variations in the first four ICMTA modes, as well as insensitivity of the other frequency domain parameters in this study may be attributed to several reasons. The “globalness” of the measure has been stated as one of the reasons for finding no effect of mental load on HRV (de Waard 1996). A study by Lee and Park (1990) showed that an increase in physical load increased HR and decreased HRV. However, in their experiment, increase in mental load reduced HRV but had no effect on HR. Hjortskov et al. (2004) stated that characteristics of the experimental stressor may be the reason for a lack of association between HRV and mental stress. Garde et al. (2002) found variations in I_{SNS} and I_{PNS} in response to a physically demanding reference computer task. They did not observe any effect of additional mental demands on these parameters. They concluded that physical demands significantly influenced I_{SNS} and I_{PNS} rather than mental demands during computer work. Other factors that affect HR include muscular fatigue and anxiety (Borghini et al. 2012).

8.5 Situation awareness

It was observed that adding the steering task to the supervisory task neither increased nor decreased situation awareness of operators. This can be due to the routineness of the steering task in a straight line in a field while seeding, as most of the participants were highly experienced. As is stated by Endsley et al. (2003), experience can cause a level of automaticity in mental processing. Cottrell and Barton (2012) also stated that for many experienced car drivers, processing some aspects of the driving task could be rather automatic, and would not draw upon their cognitive resources.

Situation awareness increased as automation support was applied to higher levels of information processing functions, however, there was a sudden drop in the highest level of automation (i.e., Action Implementation mode). This was in accordance with the statement that active involvement in the operation would increase situation awareness as opposed to acting as a supervisor of automation (Landry 2009). Action Implementation support was the only condition in which subjects were doing only a supervisory task, especially when the auto-steer was engaged.

According to the results, addition of the physical task of steering to the supervisory task of the ICMT, especially when the operators were actively involved in the task-loop (first four levels of ICMTA), caused higher attentional demand. The highest level of ICMTA (Action Implementation mode), regardless of manual steering task, substantially reduced the demand on attentional resources. A similar result was reported by Stanton and Young (2005) that a manual driving mode led to higher demand on attentional resources compared to driving with an adaptive cruise control (AAC) mode.

When examining the underlying dimensions of supply, it was found that lower ratings for Action Implementation were due to the lower ratings for arousal and concentration queries. The arousal rating shows the degree of alertness or readiness of participants for an action. Concentration of attention also explained the degree to which the subject's thoughts were brought to bear. Highly reliable automation could be one reason for having a lower rating for arousal and concentration. Out-of-the-loop taxonomy could also cause deviation in the operator's concentration on the supervisory task.

In the case of situational understanding, it was found that lower values for Manual and Action Implementation modes were due to i) information quality and ii) familiarity with the situation, respectively. Because they were responsible for all of the levels of the information processing function, operators gave lower ratings for information quality in the manual mode. However, information quantity and familiarity ratings were similar to information quality ratings. In Action Implementation mode, feeling 'out of the loop' seemed to be the cause for lower ratings of familiarity. Conversely, higher ratings for Information Analysis and Decision and Action Selection modes were mainly due to higher ratings of information quality and quantity.

Schömig and Metz (2013), on their study on driving with secondary tasks, found that the availability of corresponding cues in the environment, that allows drivers to interpret the given situation correctly, is a crucial precondition for making correct decisions. The higher quality rating in Decision and Action Selection mode in the present study verified the usefulness of providing direct solutions for system errors.

Negative association of situation awareness with mental workload was another indicator of the advantage of automation when drivers are involved in the task loop.

Keeping drivers out of the task loop in highly automated conditions decreased mental workload but also reduced situation awareness. Reaction time, number of errors, and HRV parameters showed no correlation with situation awareness. It could be inferred from this result that the lowest level of situation awareness was enough to keep a constant level of performance.

Chapter 9

CONCLUSIONS

9.1 Research findings and contributions

The goal of this study was to assess the effect of different forms of automation on behavior of operators of agricultural semi-autonomous vehicles. An experiment was performed to assess the effect of vehicle steering task automation (VSTA) and implement control and monitoring task automation (ICMTA) on mental workload and situation awareness of tractor drivers. The simultaneous measurement of mental workload, situation awareness, implement control and monitoring task performance, and HRV allowed better understanding of the interactions between operators and the automated systems they use in a tractor with an attached air seeder. The results of this study showed a greater impact of automating a supervisory task (ICMTA) rather than automating a regular manual task (VSTA).

The outcomes of the current study are summarized in the remainder of this paragraph. Increasing automation level of the supervisory task (i.e., the implement control and monitoring task) decreased mental workload, reaction time and number of errors, but automation of the steering task did not affect any of these factors. Although some parameters of HRV showed sensitivity to changes in driving conditions, in most cases, HRV was unable to differentiate mental workload levels. The widely used 0.1 Hz component of HRV, and max/min RR interval ratio presented similar results as the subjective measure did, indicating lower mental workload in highly automated supervisory task condition. When the drivers were involved in the task-loop, the 0.1 Hz component of HRV, was not sensitive enough to differentiate mental workload levels.

The results obtained from this evaluation support the hypothesis that a highly automated agricultural vehicle would reduce the situation awareness of the operator when compared with the scenario of partial automation support. Full automation of the air seeder monitoring and control task imposed out-of-the-loop consequences on operators, although this condition resulted in higher situation awareness compared to the Manual condition. The highest level of situation awareness was reported in the Decision and Action Selection mode where the system provided necessary adjustment requirements. It was also found that the engagement of the auto-steer system significantly reduced the attention required by the seeding task. The auto-steer system, however, did not affect the supply and understanding components of situation awareness.

Based on the findings of this study, the design of support systems for agricultural machines does affect the performance of the operator. Therefore, in the design of highly automated agricultural vehicles, if the presence of the operator is to be maintained, the right amount of workload must be assigned to ensure safe and efficient operation. Based on the results, automation of the Decision And Action Selection mode is a practical solution to ensure the preservation of high levels of situation awareness while operators are experiencing medium levels of mental workload. Automation of this function, in which operators were involved in the task loop, resulted in the highest level of situation awareness for medium levels of mental workload in the study. These conditions were enough to carry out the given tasks properly.

9.2 Caveats and future research directions

The results from this research provide some evidence that there are benefits to the use of automation in agricultural machines, however, caution should be exercised before making broad generalizations as there were several limitations in this study:

1. The results were obtained in a simulated environment. Although simulators offer many benefits regarding experimental control and cost, unless the results are verified with real world practice, there is some uncertainty in their application.
2. From results and observations, there was variability among participants with respect to their skills and confidence in completing the trials. Furthermore, participants were from a young generation and could not represent all of the age ranges from the population of agricultural machinery operators.
3. The air seeder display used in this study was a customized design based on previous research and was intended to represent a generalized display configuration. Although some of the participants found the information display similar to the displays they had worked with - so it was easier for them to adjust – it was unfamiliar for a few participants at first glance.
4. The training session in this study was short compared to the amount of time that an operator needs to become accustomed to a new system. It may take a few hours to a few days for an operator to acquire the complete skill of operating a new air seeder system. Similarly, driving blocks were much shorter than real world operations. Driving blocks in this study were 12 min while in real operation, operations may take a few hours before the operator takes a break.

5. Time of day for an individual’s trial varied by his/her availability. Performance of individuals may be different in different times of a day. According to Folkard (1979) “subjects engage in more maintenance processing based on the physical characteristics of the items in the morning, but more elaborative processing based on the items' meanings in the evening.”
6. There are substantial differences between tractor-air seeder systems and other machine systems used in production agriculture in terms of task type and difficulty. Furthermore, the environment may be completely different at the time of using each machine. For example, in seeding season, the field is the color of soil, but at the time of spraying, it would be green. Different impacts of a colored environment on work performance have been discussed in Jalil et al. (2012).
7. Given that the working scenario of the tractor air-seeder system in this study was unique and only pertinent to off-road operations, the results of this study cannot be considered relevant to the automation of on-road driving tasks.
8. In this study, four human factors measures were assessed simultaneously, three of which needed direct input from participants. Although using measures simultaneously has been customary in many human factors studies, based on my observations, this may distract subjects from performing the driving task naturally.

Chapter 10

FUTUTRE DIRECTIONS

Considering the findings of the study and aforementioned caveats, there are many possibilities for future work. The first suggestion is redoing the experiment discussed in this study, with longer driving periods in the simulated environment. Having longer training periods would allow the participants to better adjust to the experimental condition. Furthermore, a different experimental design would be possible to minimize the training effect. The next suggestion is to consider several agricultural machines in the experiments. After that stage, it would be useful to move from simulator studies to research using the actual machines in the field setting. It may be useful to use a commercially-available information display. Performing experiments with mental workload and situation awareness measures different from the ones used in this study also could be beneficial to determine sensitivity of the various measures. With respect to the participants, different age ranges could be considered to represent the real population of operators.

REFERENCES

Annett, J. (2003). Hierarchical task analysis. In *Handbook of cognitive task design*,Ed., E. Hollnagel, 17-35. Mahwah, NJ, US: Lawrence Erlbaum Associates.

Bahner, J. E., A. D. Hüper, and D. Manzey. 2008. Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies 66*(9): 688-699. doi: 10.1016/j.ijhcs.2008.06.001

Baldwin, C. L., F. G. Freeman, and J. T. Coyne. 2004. Mental Workload as a Function of Road Type and Visibility: Comparison of Neurophysiological, Behavioral, and Subjective Indices. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 48*(19): 2309-2313. doi: 10.1177/154193120404801922

Balfe, N., J. R. Wilson, S. Sharples, and T. Clarke. 2011. Development of design principles for automated systems in transport control. *Ergonomics 55*(1): 37-54. doi: 10.1080/00140139.2011.636456

Bedny, G. Z., and S. R. Harris. 2008. "Working sphere/engagement" and the concept of task in activity theory. *Interacting with Computers 20*(2): 251-255. doi: 10.1016/j.intcom.2007.07.005

Billings, C., E. 1991. *Human-Centered Aircraft Automation: A Concept and Guidelines*. 110381. Moffett Field, CA, USA: NASA Ames Research Centre.

Billings, C., E. 1996. *Aviation Automation: The search for a human-centered approach*. Mahwah: Lawrence Erlbaum Associates.

Borghini, G., L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni. 2012. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*. doi: http://dx.doi.org/10.1016/j.neubiorev.2012.10.003

Brookhuis, K. A., and D. de Waard. 2010. Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention 42*(3): 898-903. doi: http://dx.doi.org/10.1016/j.aap.2009.06.001

Brookhuis, K. A., C. J. G. van Driel, T. Hof, B. van Arem, and M. Hoedemaeker. 2009. Driving with a congestion assistant; mental workload and acceptance. *Applied Ergonomics 40*(6): 1019-1025. doi: http://dx.doi.org/10.1016/j.apergo.2008.06.010

Bye, A., E. Hollnagel, and T. S. Brendeford. 1999. Human-machine function allocation: a functional modelling approach. *Reliability Engineering & System Safety 64*(2): 291-300. doi: 10.1016/s0951-8320(98)00069-6

Caggiano, D., and R. Parasuraman. 2004. The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review 11*(5): 932-937. doi: 10.3758/BF03196724

Cain, B. 2007. A review of the mental workload literature: Defence Research and Development Canada (DRDC).

Calhoun, G. L., H. A. Ruff, M. H. Draper, and E. J. Wright. 2011. Automation-Level Transference Effects in Simulated Multiple Unmanned Aerial Vehicle Control.

Journal of Cognitive Engineering and Decision Making 5(1): 55-82. doi: 10.1177/1555343411399069

Cantin, V., M. Lavallière, M. Simoneau, and N. Teasdale. 2009. Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention 41*(4): 763-771. doi: 10.1016/j.aap.2009.03.019

Clamann, M. P., M. C. Wright, and D. B. Kaber. 2002. Comparison of Performance Effects of Adaptive Automation Applied to Various Stages of Human-Machine System Information Processing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 342-346.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates. Retrieved from http://books.google.ca/books?id=cIJH0lR33bgC

Cottrell, N. D., and B. K. Barton. 2012. The impact of artificial vehicle sounds for pedestrians on driver stress. *Ergonomics 55*(12): 1476-1486. doi: 10.1080/00140139.2012.724452

Crystal, A., and B. Ellington. 2004. Task analysis and human-computer interaction: approaches, techniques, and levels of analysis. In *Proceedings of the Tenth Americas Conference on Information Systems*, , 1-9. New York, NY, USA.

Cummings, M. L. 2004. Automation Bias in Intelligent Time Critical Decision Support Systems. In AIAA Intelligent Systems. Chicago, IL, USA.

de Brunélis, T. C., P. Le Blaye, P. Bonnet, and N. Maille. 2008. Distant mission management and dynamic allocation of functions. In Humans Operating Unmanned Systems (HUMOUS) conference, . Brest, France. Septtember 3-4.

de Tjerk, E. G., F. R. A. Henryk, and M. A. Neerincx. 2010. Adaptive Automation Based on an Object-Oriented Task Model: Implementation and Evaluation in a Realistic C2 Environment. *Journal of Cognitive Engineering and Decision Making 4*(2): 152-182. doi: 10.1518/155534310x522851

de Waard, D. 1996. *The measurement of drivers' mental workload.* PhD Thesis, Haren, Netherlands: University of Groningen.

de Winter, J. C. F., and D. Dodou. 2014. Why the Fitts list has persisted throughout the history of function allocation. *Cognition, Technology & Work 16*(1): 1-11. doi: 10.1007/s10111-011-0188-1

Dearden, A., M. Harrison, and P. Wright. 2000. Allocation of function: scenarios, context and the economics of effort. *International Journal of Human-Computer Studies 52*(2): 289-318. doi: 10.1006/ijhc.1999.0290

Dekker, S. 2004. On the other side of promise. What should we automate today? In *Human factors in the design of civil aircraft,* Ed., V. D. H. John A. Wise, Daniel J. Garland, Aldershot, UK: Ahsgate.

Desai, T. P. 1993. Stress and Mental Workload: A Study in an Industrial Organisation. *Indian Journal of Industrial Relations 28*(3): 258-273.

Dey, A. K., and D. D. Mann. 2010. A complete task analysis to measure the workload associated with operating an agricultural sprayer equipped with a navigation device. *Applied Ergonomics 41*(1): 146-149. doi: http://dx.doi.org/10.1016/j.apergo.2009.06.005

Dey, A. K., and D. D. Mann. 2011. Mental Workload Associated with Operating an Agricultural Sprayer: An Empirical Approach. *Journal of Agricultural Safety and Health 17*(2): 91-110.

Di Marco, L. Y., R. Sottile, and L. Chiari. 2010. Time-frequency analysis of cardio-respiratory response to mental task execution. In *Computing in Cardiology, 2010,* 753-756. 26-29 Sept. 2010.

Di Stasi, L. L., V. Álvarez-Valbuena, J. J. Cañas, A. Maldonado, A. Catena, A. Antolí, and A. Candido. 2009. Risk behaviour and mental workload: Multimodal assessment techniques applied to motorbike riding simulation. *Transportation Research Part F: Traffic Psychology and Behaviour 12*(5): 361-370. doi: 10.1016/j.trf.2009.02.004

Di Stasi, L. L., A. Antolí, and J. J. Cañas. 2011. Evaluating mental workload while interacting with computer-generated artificial environments. *Entertainment Computing*. doi: 10.1016/j.entcom.2011.03.005

Durso, F. T., and A. L. Alexander. 2009. Managing workload, performance, and situation awareness in aviation systems. In *Human Factors in Aviation (2nd Edition),* Eds., E. Salas & D. Maurino, 217-247. Burlington, MA, USA: Academic Press.

Durso, F. T., M. K. Bleckley, and A. R. Dattel. 2006. Does Situation Awareness Add to the Validity of Cognitive Tests? *Human Factors: The Journal of the Human*

Factors and Ergonomics Society 48(4): 721-733. doi: 10.1518/001872006779166316

Durso, F. T., C. A. Hackworth, T. R. Truitt, J. Crutchfield, D. Nikolic, and C. A. Manning. 1998. Situation awareness as a predictor of performance for en route air traffic controllers. *Air Traffic Control Quarterly 6*(1): 1-20. doi: citeulike-article-id:975174

Durso, F. T., and A. Sethumadhavan. 2008. Situation Awareness: Understanding Dynamic Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society 50*(3): 442-448. doi: 10.1518/001872008x288448

Edan, Y., S. Han, and N. Kondo. 2009. Automation in Agriculture. In *Springer Handbook of Automation,* Ed., S. Y. Nof, 1095-1128. Springer Berlin Heidelberg.

Endsley, M. R. 1995. Measurement of situation awareness in dynamic-systems. *Human Factors 37*(1): 65-84. doi: 10.1518/001872095779049499

Endsley, M. R. 1996. Automation and Situation Awareness. In *Automation and human performance: Theory and applications,* Eds., R. Parasuraman & M. Mouloua, 163-181. Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. R. 1999. Situation Awareness In Aviation Systems. In *Handbook of Aviation Human Factors,* Eds., D. J. Garland, J. A. Wise & V. D. Hopkin, 257-276. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Endsley, M. R. 2000. Direct measurement of situation awareness:validity and use of SAGAT. In *Situation awareness analysis and measurement,* Eds., M. R. Endsley & D. J. Garland, 131-157. Lawrence Erlbaum Associates.

Endsley, M. R. 2013. Situation Awareness. In *The Oxford Handbook of Cognitive Engineering*, Eds., J. D. Lee & A. Kirlik, Oxford, UK: Oxford University Press.

Endsley, M. R., B. Bolte, and D. G. Jones. 2003. *Designing for Situation Awareness: An Approach to User-Centered Design*. New York, NY: Taylor & Francis Inc.

Endsley, M. R., and D. J. Garland. 2000. *Situation awareness: analysis and measurement*. Lawrence Erlbaum Associates.

Endsley, M. R., S. J. Selcon, T. D. Hardiman, and D. G. Croft. 1998. A Comparative Analysis of Sagat and Sart for Evaluations of Situation Awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 42*(1): 82-86. doi: 10.1177/154193129804200119

Engström, J., E. Johansson, and J. Östlund. 2005. Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour 8*(2): 97-120. doi: http://dx.doi.org/10.1016/j.trf.2005.04.012

Etzler, L., S. Marzani, R. Montanari, and F. Tesauri. 2008. Mitigating Accident Risk in Farm Tractors. *Ergonomics in Design: The Quarterly of Human Factors Applications 16*(1): 6-13. doi: 10.1518/106480408x282737

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods 39*(2): 175-191. doi: 10.3758/BF03193146

Finomore, V., G. Matthews, T. Shaw, and J. Warm. 2009. Predicting vigilance: A fresh look at an old problem. *Ergonomics 52*(7): 791-808. doi: 10.1080/00140130802641627

Fitts, P. M. 1951. Human engineering for an effective air-navigation and traffic-control system. Washington: National Research Council, Division of Anthropology and Psychology, Committee on Aviation Psychology.

Folkard, S. 1979. Time of day and level of processing. *Memory & Cognition 7*(4): 247-252. doi: 10.3758/BF03197596

Fone, N. 2007. What happens when you fall asleep with Autotrac. http://www.fwi.co.uk/articles/07/06/2007/104197/what-happens-when-you-fall-asleep-with-autotrac.htm

Form, W. 1987. On the Degradation of Skills. *Annual Review of Sociology 13*: 29-47.

Freeman, F. G., P. J. Mikulka, M. W. Scerbo, and L. Scott. 2004. An evaluation of an adaptive automation system using a cognitive vigilance task. *Biological Psychology 67*(3): 283-297. doi: 10.1016/j.biopsycho.2004.01.002

Gabaude, C., B. Baracat, C. Jallais, M. Bonniaud, and A. Fort. 2012. Cognitive load measurement while driving. In *Proceedings of the HFES Europe Chapter Conference*, 67-80. Toulouse, France.

Galster, S. M., R. S. Bolia, and R. Parasuraman. 2002. Effects of Information Automation and Decision-Aiding Cueing on Action Implementation in a Visual Search Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 46*(3): 438-442. doi: 10.1177/154193120204600347

Garde, A., B. Laursen, A. Jørgensen, and B. Jensen. 2002. Effects of mental and physical demands on heart rate variability during computer work. *European Journal of Applied Physiology 87*(4-5): 456-461. doi: 10.1007/s00421-002-0656-7

Gempton, N., S. Skalistis, J. Furness, S. Shaikh, and D. Petrovic. 2013. Autonomous Control in Military Logistics Vehicles: Trust and Safety Analysis. In *Engineering Psychology and Cognitive Ergonomics,* Ed., D. Harris, 253-262. Springer Berlin Heidelberg.

George, C. F. P. 2003. Driving simulators in clinical practice. *Sleep Medicine Reviews 7*(4): 311-320. doi: 10.1053/smrv.2001.0233

Gordon, S. E., and R. E. Gill. 1994. Cognitive task analysis. In *Naturalistic decision making* Eds., C. E. Zsambok & G. A. Klein, 131-140.

Grote, G., S. Weik, T. Wäfler, and M. Zölch. 1995. Complementary allocation of functions in automated work systems. In *Advances in Human Factors/Ergonomics*, Eds., K. O. Yuichiro Anzai & M. Hirohiko, 989-994. Elsevier.

Gugerty, L. J. 1997. Situation Awareness During Driving: Explicit and Implicit Knowledge in Dynamic Spatial Memory. *Journal of Experimental Psychology: Applied 3*(1): 42-66.

Hahn, S. 1996. Automation of driving functions - future development, benefits and pitfalls. In *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*, 309-312. September 19-20.

Hamilton, W. I., and T. Clarke. 2005. Driver performance modelling and its practical application to railway safety. *Applied Ergonomics 36*(6): 661-670. doi: 10.1016/j.apergo.2005.07.005

Han, X.-Z., H.-J. Kim, H.-C. Moon, H.-J. Woo, J.-H. Kim, and Y.-J. Kim. 2013. Development of a Path Generation and Tracking Algorithm for a Korean Auto-guidance Tillage Tractor. *Journal of Biosystems Engineering 38, Issue ,1, 2013, pp.1-8*(1): 1-8. doi: 10.5307/JBE.2013.38.1.001

Hancock, P. A., and S. F. Scallen. 1998. Allocating functions in human–machine systems. In *Viewing psychology as a whole: The integrative science of William N. Dember.*, Eds., R. R. Hoffman, M. F. Sherrick & J. S. Warm, 509-539. Washington, DC, US: American Psychological Association.

Helton, W., and P. Russell. 2011. Working memory load and the vigilance decrement. *Experimental Brain Research 212*(3): 429-437. doi: 10.1007/s00221-011-2749-1

Hjortskov, N., D. Rissén, A. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard. 2004. The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology 92*(1-2): 84-89. doi: 10.1007/s00421-004-1055-z

Hodgkinson, G. P., and C. M. Crawshaw. 1985. Hierarchical task analysis for ergonomics research: An application of the method to the design and evaluation of sound mixing consoles. *Applied Ergonomics 16*(4): 289-299. doi: 10.1016/0003-6870(85)90094-8

Hollnagel, E., and A. Bye. 2000. Principles for modelling function allocation. *International Journal of Human-Computer Studies 52*(2): 253-265. doi: 10.1006/ijhc.1999.0288

Horiguchi, A., and T. Suetomi. 1995. A Kansei Engineering approach to a driver/vehicle system. *International Journal of Industrial Ergonomics 15*(1): 25-37. doi: 10.1016/0169-8141(94)00054-7

Hurwitz, J. B., and D. J. Wheatley. 2002. Using Driver Performance Measures to Estimate Workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1804-1808.

Hwang, S.-L., Y.-J. Yau, Y.-T. Lin, J.-H. Chen, T.-H. Huang, T.-C. Yenn, and C.-C. Hsu. 2008. Predicting work performance in nuclear power plants. *Safety Science 46*(7): 1115-1124. doi: http://dx.doi.org/10.1016/j.ssci.2007.06.005

Inagaki, T. 2003. Adaptive Automation: Sharing and Trading of Control. In *Handbook of cognitive task design*, Ed., E. Hollnagel, 147–169. London, UK: Lawrence Erlbaum Associates.

Jalil, N. A., R. M. Yunus, and N. S. Said. 2012. Environmental Colour Impact upon Human Behaviour: A Review. *Procedia-Social and Behavioral Sciences 35*: 54-62. doi: http://dx.doi.org/10.1016/j.sbspro.2012.02.062

Jensen-Urstad, K., N. Storck, F. Bouvier, M. Ericson, L. E. Lindbland, and M. Jensen-Urstad. 1997. Heart rate variability in healthy subjects is related to age and gender. *Acta Physiologica Scandinavica 160*(3): 235-241. doi: 10.1046/j.1365-201X.1997.00142.x

Johnson, A., and A. Widyanti. 2011. Cultural influences on the measurement of subjective mental workload. *Ergonomics 54*(6): 509-518. doi: 10.1080/00140139.2011.570459

Jones, D. G. 2000. Subjective measures of situation awareness. In *Situation awareness: analysis and measurement,* Eds., M. R. Endsley & D. J. Garland: Lawrence Erlbaum Associates.

Jones, D. G., and M. Endsley. 2000. Can real-time probes provide a valid measure of situation awareness? In *Human performance, situation awareness and automation: User centered design for the new millennium conference*.

Jones, D. G., and M. R. Endsley. 2004. Use of Real-Time Probes for Measuring Situation Awareness. *The International Journal of Aviation Psychology 14*(4): 343-367. doi: 10.1207/s15327108ijap1404_2

Kaber, D. B., and M. R. Endsley. 2004. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science 5*(2): 113-153. doi: 10.1080/1463922021000054335

Kaber, D. B., C. M. Perry, N. Segall, C. K. McClernon, and L. J. Prinzel. 2006. Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics 36*(5): 447-462. doi: 10.1016/j.ergon.2006.01.008

Kaber, D. B., M. C. Wright, L. J. Prinzel, and M. P. Clamann. 2005. Adaptive Automation of Human-Machine System Information-Processing Functions.

Human Factors: The Journal of the Human Factors and Ergonomics Society 47(4): 730-741. doi: 10.1518/001872005775570989

Kaminaka, M. S., and T. Fortis. 1983. Simulator for Evaluation of Dynamic Visual Fields. *Transactions of the ASABE 26*(5): 1287-1290. doi: doi: 10.13031/2013.34118

Kantowitz, B. H. 1995. Simulator Evaluation of Heavy-Vehicle Driver Workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 39*(17): 1107-1111. doi: 10.1177/154193129503901709

Kantowitz, B. H., R. J. Hanowski, and L. Tijerina. 1996. Simulator Evaluation of Heavy-Vehicle Driver Workload: II: Complex Secondary Tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 40*(18): 877-881. doi: 10.1177/154193129604001802

Käppler, W. D. 2008. *Smart driver training simulation : save money, prevent*. Berlin Springer.

Karimi, D. 2008. *Driving simulation to study the role of different sensory cues in operating an agricultural vehicle.* PhD Thesis, Winnipeg, MB, Canada: University of Manitoba. http://hdl.handle.net/1993/21202.

Karimi, D., J. Henry, and D. D. Mann. 2012. Effect of Using GPS Autosteer Guidance Systems on the Eye-Glance Behavior and Posture of Tractor Operators. *18*(4).

Karimi, D., and D. Mann. 2008. Role of motion cues in straight-line driving of an agricultural vehicle. *Biosystems Engineering 101*(3): 283-292. doi: 10.1016/j.biosystemseng.2008.09.006

Karimi, D., and D. Mann. 2009. Torque feedback on the steering wheel of agricultural vehicles. *Computers and Electronics in Agriculture 65*(1): 77-84. doi: 10.1016/j.compag.2008.07.011

Karimi, D., D. D. Mann, and J. Yan. 2011. A comparison of textual, symbolic, and pictorial presentation of information on an air-seeder display. *Australian Journal of Agricultural Engineering 2*(4): 90-95.

Kawamura, A., C. Maeda, T. Shirakawa, T. Ishida, T. Nakatsuji, and K. Himeno. 2004. Applicability of a driving simulator as a new tool for the pavement surface evaluation. In *Proceedings of the SIIV (Italian Society for Transportation) 2004 International conference*. Florence, Italy October 27-29

Kemeny, A., and F. Panerai. 2003. Evaluating perception in driving simulation experiments. *Trends in Cognitive Sciences 7*(1): 31-37. doi: 10.1016/s1364-6613(02)00011-6

Kieras, D. E. 1997. *Task Analysis and the Design of Functionality*.

Kim, G.-M., and J.-M. Woo. 2011. Determinants for Heart Rate Variability in a Normal Korean Population. *J Korean Med Sci 26*(10): 1293-1298.

Kirwan, B., and L. K. Ainsworth. 1992. *A guide to task analysis*. Philadelphia, PA, US: Taylor & Francis.

Koutsopoulos, H. N., A. Polydoropoulou, and M. Ben-Akiva. 1995. Travel simulators for data collection on driver behavior in the presence of information. *Transportation Research Part C: Emerging Technologies 3*(3): 143-159. doi: 10.1016/0968-090x(95)00003-2

Kramer, A. F., L. J. Trejo, and D. G. Humphrey. 1996. Psychophysiological measures of workload: Potential applications to adaptively automated systems. In *Automation and human performance: theory and applications,* Eds., R. Parasuraman & M. Mouloua: Lawrence Erlbaum Associates.

Kuutti, K. 1996. Activity theory as a potential framework for humann-computer interaction research. In *Context and consciousness: activity theory and human-computer interaction,* Ed., B. A. Nardi, 17-44. MIT Press.

Lafreniere, D. 1996. CUTA: a simple, practical, low-cost approach to task analysis. *interactions 3*(5): 35-39. doi: 10.1145/234757.234761

Lagu, A. V., and S. J. Landry. 2011. Roadmap for the next generation of dynamic function allocation theories and strategies. *Human Factors and Ergonomics in Manufacturing & Service Industries 21*(1): 14-28. doi: 10.1002/hfm.20209

Landry, S. 2009. Flight Deck Automation. In *Springer Handbook of Automation,* Ed., S. Y. Nof, 1215-1239. Springer Berlin Heidelberg.

Lang, T., T. Göres, D. Jünemann, M. Vollrath, J. Werneke, and A. K. Huemer. 2009. Analysis of Human Factors on agricultural machines. *Landtechnik 64*(1): 3.

Laux, L., and C. Plott. 2007. Using operator workload data to inform human reliability analyses. In *Human Factors and Power Plants and HPRCT 13th Annual Meeting, 2007 IEEE 8th,* 309-313. August 26-31.

Lee, D. H., and K. S. Park. 1990. Multivariate analysis of mental and physical load components in sinus arrhythmia scores. *Ergonomics 33*(1): 35-47. doi: 10.1080/00140139008927092

Li, P., B. Wu, Y. Zhang, Z. Yao, W. Huang, and X. Zhang. 2013. Changes in heart rate variability during manual controlled rendezvous and docking with task complexity. In *Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management. Human Body Modeling and Ergonomics,* Ed., V. Duffy, 86-92. Springer Berlin Heidelberg.

Limbourg, Q., C. Pribeanu, and J. Vanderdonckt. 2001. Towards Uniformed Task Models in a Model-Based Approach. In *Interactive Systems: Design, Specification, and Verification,* Ed., C. Johnson, 164-182. Springer Berlin Heidelberg.

Lunenfeld, H. 1989. Human factor considerations of motorist navigation and information systems. In *Vehicle Navigation and Information Systems Conference*, 35-42. September 11-13.

Luximon, A., and R. S. Goonetilleke. 2001. Simplified subjective workload assessment technique. *Ergonomics 44*(3): 229-243. doi: 10.1080/00140130010000901

Ma, R., and D. B. Kaber. 2005. Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics 35*(10): 939-953. doi: 10.1016/j.ergon.2005.04.002

Manzey, D., J. Reichenbach, and L. Onnasch. 2008. Performance Consequences of Automated Aids in Supervisory Control: The Impact of Function Allocation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 52*(4): 297-301. doi: 10.1177/154193120805200421

Marzani, S., F. Tesauri, L. Minin, R. Montanari, and C. Calefato. 2009. Designing a Control and Visualization System for Off-Highway Machinery According to the Adaptive Automation Paradigm. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience,* Eds., D. Schmorrow, I. Estabrooke & M. Grootjen, 42-50. Springer Berlin / Heidelberg.

Masalonis, A. J. 2003. Effects of training operators on situation-specific automation reliability. In *IEEE International Conference on systems, Man and Cybernetics*, 1595-1599. October 5-8.

McClernon, C. K., D. B. Kaber, C. M. Perry, and N. Segall. 2006. Towards a Sensitive Measure of Situation Awareness in Adaptively Automated Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50*(3): 275-279. doi: 10.1177/154193120605000314

McGarry, K., E. Rovira, and R. Parasuraman. 2003. Effects of Task Duration and Type of Automation Support on Human Performance and Stress in a Simulated Battlefield Engagement Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 47*(3): 548-552. doi: 10.1177/154193120304700362

Mehler, B., B. Reimer, and Y. Wang. 2011. A comparison of heart rate and heart rate variability indices in distinguishing single task driving and driving under secondary cognitive workload. In *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*

590-597. Lake Tahoe, California, USA.

Meshkati, N. 1988. Heart Rate Variability and Mental Workload Assessment. In *Advances in Psychology,* Eds., A. H. Peter & M. Najmedin, 101-115. Elsevier Science.

Moray, N. 2003. Monitoring, complacency, scepticism and eutactic behaviour. *International Journal of Industrial Ergonomics 31*(3): 175-178. doi: 10.1016/s0169-8141(02)00194-4

Mulder, L. J. M., D. de Waard, and K. A. Brookhuis. 2004. Estimating mental effort using heart rate and heart rate variability. In *Handbook of Human Factors and Ergonomics Methods*: CRC Press.

Nayem, M., N. Begum, and S. Ferdousi. 2013. Time domain measures of heart rate variability to assess autonomic dysfunction in irritable bowel syndrome. *Journal of Bangladesh Society of Physiologist 7*(2): 60-65. doi: http://dx.doi.org/10.3329/jbsp.v7i2.14451

Nickel, P., and F. Nachreiner. 2003. Sensitivity and diagnosticity of the 0.1-Hz component of heart rate variability as an indicator of mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society 45*(4): 575-590. doi: 10.1518/hfes.45.4.575.27094

Norman, D. A. 1990. The 'Problem' with Automation: Inappropriate Feedback and Interaction, not 'Over-Automation'. *Philosophical Transactions of The Royal Society B: Biological Sciences 327*(1241): 585-593. doi: 10.1098/rstb.1990.0101

Onnasch, L., C. D. Wickens, H. Li, and D. Manzey. 2014. Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis.

Human Factors: The Journal of the Human Factors and Ergonomics Society 56(3): 476-488. doi: 10.1177/0018720813501549

Parasuraman, R. 1986. Vigilance, monitoring, and search. In *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance.*, Eds., K. R. Boff, L. Kaufman & J. P. Thomas, 1-39. Oxford, England: John Wiley & Sons.

Parasuraman, R., M. Barnes, and K. Cosenzo. 2007. Adaptive Automation for Human-Robot Teaming in Future Command and Control Systems. *The International C2 Journal 1*(2): 25.

Parasuraman, R., R. Molloy, and I. Singh. 1993. Performance Consequences of Automation-Induced 'Complacency'. *International Journal of Aviation Psychology 3*(1): 1-23. doi: 10.1207/s15327108ijap0301_1

Parasuraman, R., T. B. Sheridan, and C. D. Wickens. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics 30*(3): 286-297. doi: 10.1109/3468.844354

Parasuraman, R., and C. D. Wickens. 2008. Humans: Still Vital After All These Years of Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society 50*(3): 511-520. doi: 10.1518/001872008x312198

Parnas, D. L., J. van Schouwen, and S. P. Kwan. 1990. Evaluation of safety-critical software. *Communications of ACM 33*(6): 636-648. doi: 10.1145/78973.78974

Pattyn, N., X. Neyt, D. Henderickx, and E. Soetens. 2008. Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior 93*(1-2): 369-378. doi: 10.1016/j.physbeh.2007.09.016

Pauzié, A. 2008a. Evaluating driver mental workload using the driving activity load index(DALI). In European conference on human centred design for intelligent transport systems. Lyon, France. April 3-4.

Pauzié, A. 2008b. A method to assess the driver mental workload: The driving activity load index (DALI). *Intelligent Transport Systems, IET 2*(4): 315-322. doi: 10.1049/iet-its:20080023

Pauzié, A., J. Manzano, and N. Dapzol. 2007. Driver's behavior and workload assessment for new in-vehicle technologies design. In *In Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 572-580. Stevenson, Washington, USA. July.

Pauzié, A., A. Sarpedon, and G. Saulnier. 1995. Ergonomic evaluation of a prototype guidance system in an urban area. Discussion about methodologies and data collection tools. In *Proceedings of Vehicle Navigation and Information Systems Conference*, 390-396.

Petzoldt, T., N. Bär, C. Ihle, and J. F. Krems. 2011. Learning effects in the lane change task (LCT)—Evidence from two experimental studies. *Transportation Research Part F: Traffic Psychology and Behaviour 14*(1): 1-12. doi: http://dx.doi.org/10.1016/j.trf.2010.09.001

Preston, T. A. 1979. Simulation in ergonomics, in particular in agriculture and forestry. *Agricultural Systems 4*(2): 115-118. doi: 10.1016/0308-521x(79)90023-4

Prinzel, L. J., F. G. Freeman, M. W. Scerbo, P. J. Mikulka, and A. T. Pope. 2003a. Effects of a Psychophysiological System for Adaptive Automation on

Performance, Workload, and the Event-Related Potential P300 Component. *Human Factors: The Journal of the Human Factors and Ergonomics Society 45*(4): 601-614. doi: 10.1518/hfes.45.4.601.27092

Prinzel, L. J., R. Parasuraman, F. G. Freeman, M. W. Scerbo, P. J. Mikulka, and A. T. Pope. 2003b. *Three experiments examining the use of electroencephalogram, event-related potentials and heart-rate variability for real time human-centered adaptive automation design*. NASA/TP-2003-212442. National Aeronautics and Space Administration, Johnson Space Center.

Rajendra, A. U., J. K. Paul, N. Kannathal, C. M. Lim, and J. S. Suri. 2006. Heart rate variability: a review. *Medical and Biological Engineering and Computing 44*(12): 1031-1051. doi: 10.1007/s11517-006-0119-0

Rakauskas, M. E., L. J. Gugerty, and N. J. Ward. 2004. Effects of naturalistic cell phone conversations on driving performance. *Journal of Safety Research 35*(4): 453-464. doi: 10.1016/j.jsr.2004.06.003

Ramaekers, D., H. Ector, A. E. Aubert, A. Rubens, and F. Van de Werf. 1998. *Heart rate variability and heart rate in healthy volunteers. Is the female autonomic nervous system cardioprotective?* (Vol. 19).

Ramon, C., A. Clarion, C. Gehin, C. Petit, C. Collet, and A. Dittmar. 2008. An integrated platform to assess driver's physiological and functional states. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 506-509. 20-25 Aug. 2008.

Redding, R. E. 1990. Taking cognitive task analysis into the field: bridging the gap from research to application. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, 1304-1308.

Richardson, J., T. C. Ormerod, and A. Shepherd. 1998. The role of task analysis in capturing requirements for interface design. *Interacting with Computers 9*(4): 367-384. doi: 10.1016/s0953-5438(97)00036-2

Riley, V. 1989. A general model of mixed-initiative human-machine systems. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, 124-128.

Ryan, S. M., A. L. Goldberger, S. M. Pincus, J. Mietus, and L. A. Lipsitz. 1994. Gender- and age-related differences in heart rate dynamics: Are women more complex than men? *Journal of the American College of Cardiology 24*(7): 1700-1707. doi: http://dx.doi.org/10.1016/0735-1097(94)90177-5

Ryder, J., and R. Redding. 1993. Integrating cognitive task analysis into instructional systems development. *Educational Technology Research and Development 41*(2): 75-96. doi: 10.1007/bf02297312

Salmon, P. M., N. A. Stanton, G. H. Walker, D. Jenkins, D. Ladva, L. Rafferty, and M. Young. 2009. Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics 39*(3): 490-500. doi: 10.1016/j.ergon.2008.10.010

Sanchez, J., C. J. Krumbholz, and M. O. Silva. 2010. *Controller-automation interaction in NextGen: a new paradigm*. https://www.mitre.org/publications/technical-papers/controllerautomation-interaction-in-nextgen-a-new-paradigm

Sato, N., and S. Miyake. 2004. Cardiovascular reactivity to mental stress: relationship with menstrual cycle and gender. *Journal of PHYSIOLOGICAL ANTHROPOLOGY and Applied Human Science 23*(6): 215-223.

Scallen, S. F., and P. A. Hancock. 2001. Implementing Adaptive Function Allocation. *The International Journal of Aviation Psychology 11*(2): 197-221. doi: 10.1207/s15327108ijap1102_05

Scerbo, M. W. 1996. Theoretical perspectives on adaptive automation. In *Automation and human performance: theory and applications*, Eds., R. Parasuraman & M. Mouloua, 37-63. Lawrence Erlbaum Associates.

Schömig, N., and B. Metz. 2013. Three levels of situation awareness in driving with secondary tasks. *Safety Science 56*(0): 44-51. doi: http://dx.doi.org/10.1016/j.ssci.2012.05.029

Schraagen, J. M., S. F. Chipman, V. L. Shalin, and V. Shalin. 2000. *Cognitive task analysis*. L. Erlbaum Associates.

Schrauf, M., A. Sonnleitner, M. Simon, and W. E. Kincses. 2011. EEG Alpha Spindles as Indicators for Prolonged Brake Reaction Time During Auditory Secondary Tasks in a Real Road Driving Study. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 217-221. September 1, 2011.

Sethumadhavan, A. 2009. *Effects of Levels of Automation on Air Traffic Controller Situation Awareness and Performance.* PhD Thesis, Lubbock, Texas, US: Texas Tech University

Shepherd, A. 1998. HTA as a framework for task analysis. *Ergonomics 41*(11): 1537-1552. doi: 10.1080/001401398186063

Sheridan, T. B. 1991. *Human factors of driver-vehicle interaction in the IVHS environment*. DOT HS 807–837. Washington, DC, USA: U.S. Dept. of Transportation, National Highway Traffic Safety Administration.

Sheridan, T. B. 2002. *Humans and Automation: System Design and Research Issues*. John Wiley.

Sheridan, T. B., and W. L. Verplank. 1978. *Human and Computer Control of Undersea Teleoperators*. MIT Man-Machine Systems Laboratory

Siegler, I., G. Reymond, A. Kemeny, and A. Berthoz. 2001. Sensorimotor integration in a driving simulator: contributions of motion cueing in elementary driving tasks. In Driving Simulator Conference. Sophia Antipolis, France.

Soliman, A. M., and E. K. Mathna. 2009. Metacognitive strategy training improves driving situation awareness. *SOCIAL BEHAVIOR AND PERSONALITY 37*(9): 1161-1169. doi: 10.2224/sbp.2009.37.9.1161

Sonnenwald, D. H., K. L. Maglaughlin, and M. C. Whitton. 2004. Designing to support situation awareness across distances: an example from a scientific collaboratory. *Information Processing & Management 40*(6): 989-1011. doi: 10.1016/j.ipm.2003.10.002

Spinuzzi, C. I. 1997. Context and consciousness: Activity theory and human-computer interaction. *Computers and Composition 14*(2): 301-304. doi: 10.1016/s8755-4615(97)90030-x

Stanton, N., and P. Marsden. 1996. From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science 24*(1): 35-49. doi: http://dx.doi.org/10.1016/S0925-7535(96)00067-7

Stanton, N., and U. Middlesex. 2006. Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics 37*(1): 55-79. doi: 10.1016/j.apergo.2005.06.003

Stanton, N., M. Young, G. Walker, H. Turner, and S. Randle. 2001. Automating the Driver's Control Tasks. *International Journal of Cognitive Ergonomics 5*(3): 221-236. doi: 10.1207/s15327566ijce0503_5

Stanton, N., and M. S. Young. 2005. Driver behaviour with adaptive cruise control. *Ergonomics 48*(10): 1294-1313. doi: 10.1080/00140130500252990

Steele, M., and R. B. Gillespie. 2001. Shared control between human and machine: using a haptic steering wheel to aid in land vehicle guidance. In Human Factors and Ergonomics Society 45th Annual Meeting. Minneapolis, MN, USA. October 8-12.

Stentz, A., C. Dima, C. Wellington, H. Herman, and D. Stager. 2002. A System for Semi-Autonomous Tractor Operations. *Autonomous Robots 13*(1): 87-104. doi: 10.1023/A:1015634322857

Strybel, T. Z., K.-P. L. Vu, J. P. Dwyer, J. Kraft, T. K. Ngo, V. Chambers, and F. P. Garcia. 2007. Predicting perceived situation awareness of low altitude aircraft in terminal airspace using probe questions. In Proceedings of the 12th international conference on Human-computer interaction: interaction design and usability, 939-948. Beijing, China.

Swain, A. D. 1990. Human reliability analysis: Need, status, trends and limitations. *Reliability Engineering & System Safety 29*(3): 301-313. doi: 10.1016/0951-8320(90)90013-d

Taelman, J., S. Vandeput, A. Spaepen, and S. Van Huffel. 2008. Influence of Mental Stress on Heart Rate and Heart Rate Variability. In *IFMBE Proceedings*, 1366-1369.

Takahashi, A. M., A. Porta, R. Melo, R. Quitério, E. da Silva, A. Borghi-Silva, . . . A. Catai. 2012. Aging reduces complexity of heart rate variability assessed by conditional entropy and symbolic analysis. *Internal and Emergency Medicine 7*(3): 229-235. doi: 10.1007/s11739-011-0512-z

Thayer, J. F., J. J. Sollers Iii, D. M. Labiner, M. Weinand, A. M. Herring, R. D. Lane, and G. L. Ahern. 2009. Age-related differences in prefrontal control of heart rate in humans: A pharmacological blockade study. *International Journal of Psychophysiology 72*(1): 81-88. doi: http://dx.doi.org/10.1016/j.ijpsycho.2008.04.007

Veltman, J. A., and A. W. K. Gaillard. 1993. Indices of Mental Workload in a Complex Task Environment. *Neuropsychobiology 28*(1-2): 72-75.

Vidulich, M. A. 2002. Mental workload and situation awareness: Essential concepts for aviation psychology practice. In *Principles and practice of aviation psychology,* Eds., P. S. Tsang & M. A. Vidulich, 115-146. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Vidulich, M. A., and P. S. Tsang. 2012. Mental workload and situation awareness. In *Handbook of Human Factors and Ergonomics,* Ed., G. Salvendy, 243-273. Hoboken: John Wiley & Sons.

Vidulich, M. A., and P. S. Tsang. 2014. The Confluence of Situation Awareness and Mental Workload for Adaptable Human–Machine Systems. *Journal of Cognitive Engineering and Decision Making*. doi: 10.1177/1555343414554805

Walker, G. H., N. A. Stanton, and M. S. Young. 2006. The ironies of vehicle feedback in car design. *Ergonomics 49*(2): 161-179. doi: 10.1080/00140130500448085

Walker, G. H., N. A. Stanton, and M. S. Young. 2008. Feedback and driver situation awareness (SA): A comparison of SA measures and contexts. *Transportation Research Part F: Traffic Psychology and Behaviour 11*(4): 282-299. doi: 10.1016/j.trf.2008.01.003

Wang, X., J. F. Thayer, F. Treiber, and H. Snieder. 2005. Ethnic Differences and Heritability of Heart Rate Variability in African- and European American Youth. *The American Journal of Cardiology 96*(8): 1166-1172. doi: http://dx.doi.org/10.1016/j.amjcard.2005.06.050

Wei, J., and G. Salvendy. 2004. The cognitive task analysis methods for job and task design: review and reappraisal. *Behaviour & Information Technology 23*(4): 273-299. doi: 10.1080/01449290410001673036

Weir, D. H. 2010. Application of a driving simulator to the development of in-vehicle human-machine-interfaces. *IATSS Research 34*(1): 16-21. doi: 10.1016/j.iatssr.2010.06.005

Wheeler, W., J. Lee, M. Raby, R. Kinghom, A. Bittner, and M. McCallum. 1996. *Development of Human Factors Guidelines for Advanced Traveler Information Systems (ATIS) and Commercial Vehicle Operations (CVO): Display Formats and Commercial Vehicle Operator (CVO) Workload*. FHWA-RD-95-176. McLean, VA, USA: Federal Highway Administration.

Wickens, C. D. 1996. Situation awareness: impact of automation and display technology. *Situation awareness: limitations and enhancement in the aviation environment*: k2.

Wickens, C. D. 2008. Situation Awareness: Review of Mica Endsley's 1995 Articles on Situation Awareness Theory and Measurement. *Human Factors: The Journal of the Human Factors and Ergonomics Society 50*(3): 397-403. doi: 10.1518/001872008x288420

Wickens, C. D., and J. G. Hollands. 2000. *Engineering psychology and human performance Third Edition*. Electronic Publishing Services Inc.

Wickens, C. D., H. Li, A. Santamaria, A. Sebok, and N. B. Sarter. 2010. Stages and Levels of Automation: An Integrated Meta-analysis. *Proceedings of the Human*

Factors and Ergonomics Society Annual Meeting 54(4): 389-393. doi: 10.1177/154193121005400425

Wickens, C. D., A. S. Mavor, J. McGee, and N. R. C. P. o. H. F. i. A. T. C. Automation. 1997. *Flight to the future: human factors in air traffic control*. National Academy Press.

Wierwille, W. W., and F. T. Eggemeier. 1993. Recommendations for Mental Workload Measurement in a Test and Evaluation Environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society 35*(2): 263-281. doi: 10.1177/001872089303500205

Young, M. S., and N. A. Stanton. 1997. Automotive automation: Investigating the impact on drivers' mental workload. *International Journal of Cognitive Ergonomics 1*(4): 325-336.

Yung-Tsan, J., Y. Tzu-Chung, L. Chiuhsiang Joe, Y. Chih-Wei, and L. Shiau-Feng. 2009. Evaluation of mental workload in automation design for a main control room task. In *International Conference on Networking, Sensing and Control*, 313-317. 26-29 March 2009.

Zhang, Q., J. F. Reid, and N. Noguchi. 1999. Agricultural vehicle navigation using multiple guidance sensors. In *The International Conference on Field and Service Robotics.*

Zhang, Y., S. Jin, M. Garner, P. Mosaly, and D. Kaber. 2009 The effects of aging and cognitive stress disposition on driver situation awareness and performance in hazardous conditions. In *Proceedings of the IEA 2009 17th World Congress on Ergonomics*. Beijing, China.

APPENDIX A

UNIVERSITY OF MANITOBA | Research Ethics and Compliance

Office of the Vice-President (Research and International)

Human Ethics
208-194 Dafoe Road
Winnipeg, MB
Canada R3T 2N2
Phone +204-474-8880
Fax +204-269-7173

APPROVAL CERTIFICATE

August 9, 2012

TO: **Behzad Bashiri** (Advisor D. Mann)
Principal Investigator

FROM: **Stan Straw, Chair** [redacted]
Education/Nursing Research Ethics Board (ENREB)

Re: **Protocol #E2012:066**
"Task Analysis and Function Allocation of Agricultural Semi-autonomous Vehicles from a Human Factors Perspective"

Please be advised that your above-referenced protocol has received human ethics approval by the **Education/Nursing Research Ethics Board**, which is organized and operates according to the Tri-Council Policy Statement (2). **This approval is valid for one year only**.

Any significant changes of the protocol and/or informed consent form should be reported to the Human Ethics Secretariat in advance of implementation of such changes.

Please note:

- **If you have funds pending human ethics approval, the auditor requires that you submit a copy of this Approval Certificate to the Office of Research Services, fax 261-0325 - please include the name of the funding agency and your UM Project number. This must be faxed before your account can be accessed.**

- **if you have received multi-year funding for this research, responsibility lies with you to apply for and obtain Renewal Approval at the expiry of the initial one-year approval; otherwise the account will be locked.**

The Research Quality Management Office may request to review research documentation from this project to demonstrate compliance with this approved protocol and the University of Manitoba *Ethics of Research Involving Humans*.

The Research Ethics Board requests a final report for your study (available at: http://umanitoba.ca/research/orec/ethics/human_ethics_REB_forms_guidelines.html**) in order to be in compliance with Tri-Council Guidelines.**

APPENDIX B

Consent Form for tractor drivers

Task analysis and function allocation of agricultural semi-autonomous vehicles from a human factors perspective

Investigator: Behzad Bashiri, PhD student, and Danny D. Mann, Professor, Department of Biosystems Engineering, University of Manitoba.

Research Objective

In recent years many automated systems have been introduced in agricultural vehicles to increase their productivity. We may be tempted to think that these automated systems will replace operators because they are better able to complete the task. With current technology, however, a more appropriate view is that the system design should be based on operator and machine collaboration. In last two decades researchers have produced lists of automation levels to better promote human - machine collaboration. Automation levels include some intermediate levels between manual and full automation. In intermediate levels of automation, if the machine does not provide adequate feedback to the human operator, human errors will arise when the situations exceed the capabilities of the automatic equipment.

Automated vehicle navigation systems, as the most important automated system introduced in agricultural vehicles, made it possible to reduce guidance errors by offering a steering assist system, an automatic steering system, and route planning. Today, farmers are interacting with semi-autonomous agricultural vehicles. In such vehicles, intermediate levels of automation can be easily identified. To reduce the chance of human error, a human factors perspective is needed to ensure safe and efficient operation of these machines. In this project, we will address several problems associated with operating a semi-autonomous agricultural vehicle and then will provide some recommendations for designing such vehicles. In this research, a tractor air-seeder system will be considered as a case study.

The objectives of this research will be to determine the types and levels of automation for air-seeder control and monitoring task; and to investigate the effects of task automation on some human performance measures such as mental workload and situation awareness.

Use of a driving simulator is proposed to ensure that uncontrollable factors associated with field research can be avoided.

Research Procedure

The Principal Investigator will conduct a series of simulator experiments to compare the tractor driver's experienced workload and his/her situation awareness when exposed to different types of task automation. A driving scenario will be created that requires subjects to virtually seed a field with mentioned settings. A maximum of 10 levels of task automation will be considered. For example, in the manual setting (i.e., no task automation of any kind), the air-seeder control tasks as well as the tractor steering task will be performed manually. At the other extreme, (i.e., full task automation), the driver will perform only some supervisory tasks. The remaining levels of task automation will include a combination of manual and automated tasks. Each participant will be exposed to five automation conditions, up to 15 min each. The total time of experiments, including the amount of time needed to carry out the subjective assessment and training, is estimated two hours.

Workload will be assessed using three different methodologies. In performance based methodology, reaction time and response errors of the drivers will be collected. The simulator code has been programmed to automatically track both reaction time and response errors. For subjective assessment the Driving Activity Load Index (DALI), which is a modified version of NASA TLX for driving context, will be used. In this post-trial method, drivers will rate their experienced workload after each driving session. DALI will measure the following dimensions of workload: effort of attention, visual demand, auditory demand, temporal demand, interference and situational stress. Physiological measures of workload will be considered beside other mental workload measures for better indication of the drivers' state. Heart rate variability (HRV) has shown to be a proper measure for this purpose. For HRV measurement, a Polar heart-rate monitor (Polar S810i) consisting of a transmitter equipped with electrodes and a receiver will be used. The HRV data will be recorded during the test for later analysis.

In the case of situation awareness, only one subjective report will be collected. Situational Awareness Rating Technique (SART) is a post-trial, subjective rating technique in which subjects rate their perceived situation awareness. This multi-dimensional technique can measure various aspects of situation awareness depending on the number of dimensions. For example a three dimensional SART provides questions about complexity, variability, and instability of the situation. Immediately after each session, subjects will be asked to answer the SART questions.

Risk

All experimental procedures will be conducted using a stationary tractor-driving simulator located in the Agricultural Ergonomics Laboratory. Therefore, the risks to research subjects in this study are unlikely and minimal. They include: (1) a risk of Simulation Sickness due to the immersive nature of the simulator used in this study; (2) possible soreness of the hand and back muscles from extensive use of the steering wheel interface and turning back to monitor the seeder; and (3) potential visual strain and/or fatigue in viewing the simulation displays through projected pictures and LCD monitors. These risks are not substantially different from those associated with your everyday PC use and driving and are reversible. In the event that you indicate fatigue or discomfort during the described experiment, a rest period will be provided. If abnormal physiologic conditions persist, your participation in the experiment will be terminated.

Instruments

Task performance (i.e., reaction time, errors) will be automatically recorded by the simulator's control system while the experiment is in session. A Polar heart-rate monitor will be used to simultaneously measure heart rate variability (HRV) as an indication of experienced workload. The Polar heart-rate monitor is non-invasive and poses no risk to you.

Assurance of Confidentiality

The information in study records will be kept strictly confidential. Data will be stored securely in the Human factors and Ergonomics Lab of the Department of Biosystems Engineering. Research subjects will only be represented as a number in the test data which will not be linked to their identity. No reference will be made in oral or written reports which could link you to the study.

Availability of Research Results

Results of this experiment will be available in the form of a summary sheet six months after the date of experiment.

☐ Check the box to the left if you would like to receive a summary of the research results. Please provide your e-mail or postal address so that I can contact you when it is ready.

____

Remuneration

You will receive an honorarium of $40 for participating in the experiments.

Assurance of Voluntary Participation

Your participation in this research is voluntary. If at any time you wish to withdraw from the project, you may do so without consequence. If you decide to do so any time during the experiment, you should notify the experimenter. If you make such a request, the experiment will be stopped and the any data collected will be deleted immediately.

Human Subject Research Ethics Approval

This research has received approval by the Education/Nursing Research Ethics Board (ENREB). Any complaint regarding a procedure may be reported to either the Human Ethics Secretariat (474-7122) or the Head of Biosystems Engineering (474-9819).

If you have any questions or concerns, please contact the primary investigator, Behzad Bashiri, or his advisor, Dr. Danny Mann:

Behzad Bashiri
Department of Biosystems Engineering
University of Manitoba
Winnipeg, MB R3T 5V6
Phone: (204) 474-7446
E-mail: behzad_bashiri@umanitoba.ca

Dr. Danny Mann, P.Eng.
Department of Biosystems Engineering
University of Manitoba
Winnipeg, MB R3T 5V6
Phone: (204) 474-7149
E-mail: Danny.Mann@ad.umanitoba.ca

My signature indicates that I have read and understand the above conditions. I hereby give my consent for, and agree to participate in, this research project.

Name:__________________________________ Date: _______________________

Witnessed by: ____________________________ Date: _______________________

Demographic questionnaire

We would like to know more about you and your tractor operation experience. Please answer the questions below accurately.

* First Name:
* Last Name:
* Age:
* Sex:

- o Ethnicity:
- o First Language:

  * How long have you had your driver's license?
  * How long is your tractor (or any agricultural vehicles) driving experience?
  * with which implements?

  * Do you have any experience with the University of Manitoba tractor driving simulator?
  * Are you familiar with GPS navigation systems?

Situation Awareness Rating Technique (SART)

Please answer these questions with regard to the driving situations presented in the scenario.

Instability of Situation

How changeable is the situation? Is the situation highly unstable and likely to change suddenly (high), or is it very stable and straightforward (low)?

Low … High

Complexity of Situation

How complicated is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)?

Low … High

Variability of Situation

How many variables are changing in the situation? Are there are large number of factors varying (high) or are there very few variables changing (low)?

Low … High

Arousal

How aroused are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)?

Low … High

Concentration of Attention

How much are you concentrating on the situation? Are you bringing all your thoughts to bear (high) or is your attention elsewhere (low)?

Low … High

Division of Attention

How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (high) or focussed on only one (low)?

Low High

Spare Mental Capacity

How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (high) or nothing to spare at all (low)?

Low High

Information Quantity

How much information have you gained about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)?

Low High

Information Quality

How good is the information you have gained about the situation? Is the knowledge communicated very useful (high) or is it a new situation (low)?

Low High

Familiarity with Situation

How familiar are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

Low High

DALI – Driving Activity load Index

During the experiment, you may have a different experience compared to regular tractor (or any agricultural vehicles) driving. Table 1 contains 5 factors that will help us to evaluate your experience in different driving conditions.

Table 1. Factors Description

Global attention demand	Mental (to think about, to decide) visual and auditory demand required during the test to achieve the whole activity.
Visual demand	Visual demand required during the test to achieve the whole activity.
Stress	Level of stress during the whole activity such as fatigue, insecure feeling, irritation, discouragement.
Temporal demand	Pressure and specific constraint felt due to timing demand when running the whole activity.
Interference	Disturbance of the driver' state and consequences on the driving activity when conducting the driving activity simultaneously with any other supplementary task such as phoning, using systems or radio

In the next page, please rate each factor based on constraints you felt during a driving session on a scale from 1 (low) to 20 (high).

Global Attention Demand

How do you rate the global attention required during the test with regard to what you usually feel while driving a tractor?

Low High

Visual Demand

How do you rate the visual demand required during the test with regard to what you usually feel while driving a tractor?

Low High

Stress

How do you rate the stress required during the test with regard to what you usually feel while driving a tractor?

Low High

Temporal Demand

How do you rate the pressure related to the time available to run the whole activity during the test with regard to what you usually feel while driving a tractor?

Low High

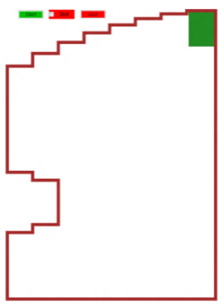
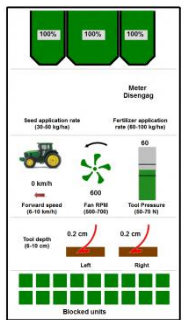
Inference

How do you rate the modifications of your driving behavior during the test with regard to what you usually feel while driving a tractor?

Low High

Post-Trial Questionnaire

1- During the experiments how much time (in percent) did you spend on supervising following items: (Total 100%)

Map/Navigation	Implement parameters	Path (visual scenery)	Implement monitoring	Other
				
----------%	--------- %	--------- %	--------- %	----- %

2- How could you find the parameters in the console?
Very Easy □ Easy □ Neutral □ Difficult □ Very Difficult □

3- How do you feel about answering identical questionnaires after each driving session?

4- If you would like to make any comments (compliments, suggestions or complaints) on the experiments, display configuration, simulator, messages on the monitor,… please provide it below:

APPENDIX C

Answers to the open-ended queries in the Post-Trial Questionnaire

3- How do you feel about answering identical questionnaires after each driving session?

#	Comments from pilot study
1	Overall, it was fine. In some situations some of the questions were not relevant (i.e. full automation condition)
2	I don’t really know if my answers are accurate because the questions are too general
3	It makes sense to get those questions answered every time to assess each aspect of the research. I found that it gave me the rest that I needed.
4	Answering first time is found a little demanding. In subsequent sessions, it was easy. I suggest using more general language in more explicit way.
5	It is good because I can answer questions based on the experience I just had for each session
6	I felt good. It helped/enabled me to answer the questions quickly.
7	A little tedious but okay otherwise
8	Becomes automatic because I know the questions already
9	Makes sense to try to evaluate same things. After a few runs I kind of know better the position of bottoms
10	Easy and was comparable

#	Comments from the main experiment
1	It was fine
2	Good, because my perspective changed with the different levels
3	Compare feeling from last experiment
4	Neutral
5	Hard to measure each by the same measure
6	Thought there was overlap in questions
7	Good, knew what to think about and monitor personally.
8	It was okay. Not many answers changed foe me though
9	Did not bother me
10	It was easier to compare each trial against each other
11	A little repetitive
12	Creates good experiment control. Might allow for more thought into the questions though if they were worded slightly different
13	They become more familiar each time to answer
14	Fine
15	good
16	It was nice because it allowed for everything to stay fresh in my mind after each trial
17	Does not really bother me
18	I found them hard to answer after a while due to fatigue and or boredom
19	Helps to realize what made each scenario easier/harder. Better compares scenarios.
20	I had no issues with them. My opinion on some aspects changed between sessions, so it makes sense to record the changes

21	I think my answers were similar for all of the questionnaires
22	They were a good way of assessing the differences and similarities between each session
23	I was not always sure if my information was accurate relative to the previous questionnaires
24	Did not matter that they were the same
25	Feel like I answered very similarly every time
26	Fine
27	I feel that it was useful as it got me to really think about what changed between tests and how it changed my state of mind
28	Fine
29	okay

4- If you would like to make any comments (compliments, suggestions or complaints) on the experiments, display configuration, simulator, messages on the monitor… please provide it below:

#	**Comments from pilot study**
1	The buttons in the console should be displaced in the same way they are arranged in the screen
2	The pop up message would be better read if are on the top of the screen (in level with the eye of the driver)
3	Overall, the experience was quite good for me. It added value to me in understanding the user interface from a different perspective.
4	Nice conductor Experiment is relatively easy, changes on the monitor can be easily detected and act before it went outside of the range I like the last one, when the message is telling the operator what to do to fix the error so I noticed the warning sign right away and could easily make adjustments. I did not have to pay much attention at all.

5	Good experiment
6	Noise got a bit annoying after a while Display was realistic and effective I found decision-making condition easiest to handle (with the exception of the Action automation mode), but I found I got bored with this test and grew unattentive.
7	Touch screen monitor for controls
8	Good simulation overall A bit tired afterwards, mostly the eyes.
9	Along with the visual display a sound alert will be useful to prompt the action

#	**Comments from the main experiment**
1	It was good
2	Second systems is the best If there are also some sounds to indicate the situation happening it should be better
3	Landscape does not change, easy to lose concentration
4	Navigation map was frustrating at times with delays from steering
5	You should ask random questions like how many bins where there to see if the person only looked at the monitor or if they also looked elsewhere, like where they were going if they were driving.
6	Visual scenery did not match with what you did while driving
7	The error message helped gave time to find the parameters before they got to too low/high of a level
8	Very realistic simulator, however, sudden and dramatic changes of variables were a little bit unrealistic. The messages were a bit distracting instead of observing the numbers but still helpful. In addition, not having the tractor visually turn around on the path was odd; it could probably be more realistic if done so.
9	Extraneous motions in the steering would help create a more realistic situation i.e.

	simulate side hills, large rocks, etc. I did not focus at all on the visual scenery, which is a very large bias in the experiment. Normally I would focus by for the most on the land ahead of me. This changes reaction time significantly. Also very hard to simulate fatigue in 5-11 min sessions whereas a normal day is three four hour sessions approximately
10	For the warning messages, I know that often monitor/gps system, etc., will manually make a sound, such as a loud beep to alert the operator (this is in real life). This did not affect me as I had auto-steer but it may help when manual steering.
11	Very well-done program and set up
12	had a decent of trouble figuring out the steering
13	When the warning messages would say for example "decrease fan speed" it was helpful, as I did not have to check what the problem was, then find the control and then fix it. Much less intensive when the warning have detail
14	I found the warning messages startling and I found it lowered by reaction time so I was already aware that the error was going to happen.
15	I could not see the mirror in the left implement very well.
16	To make it more realistic, more parameters should change at once. Very rarely do you get to fix one thing, then wait for another to change. It is a constant battle to optimize performance
17	I thought the simulator parameters for the air seeder was very easy to understand probably because I have worked with GPS air seeder equipment before.
18	Very interesting experiment! The monotonous tractor sounds are bothersome after a while though
19	Move monitor forward, reduce the need to turn head and eyes
20	The steering is glitchy
21	The steering was backwards to what I would have normally expected on some of the passes. Very good Experiment and clearly displays the usefulness of warning messages to alert the operator of a potential problem. Without warning messages, I found myself having to focus and concentrate a lot harder on the monitor.
22	Very user-friendly overall