ABSTRACT

MA, RUIQI. The Effect of In-Vehicle Automation and Reliability on Driver Situation Awareness and Trust. (Under the direction of Dr David B. Kaber).

The cognitive construct of situation awareness (SA) has not been well developed in the domain of driving. The objective of this study was to define a new transactional model of SA in various driving behaviors and activities, as influenced by automation and in-vehicle device use. Specifically, this study investigated the implications of adaptive cruise control (ACC) and cellular phone use in driving on a direct and objective measure of SA; investigate the effect of varying reliability of in-vehicle automation (navigation aids) on driver SA and trust; and assess differences in human trust in a human aid versus an automation aid in a simulated driving task.

Twenty participants drove a virtual car and performed a freeway driving task (Experiment A) as well as a suburb navigation task (Experiment B). In the freeway driving, participants were required to drive using ACC or manual control modes, and received navigation information from one of two sources: a human or in-vehicle automation aid via cell phone or separate display screen, respectively. During the navigation driving, participants were required to drive through the suburban area following all traffic signs and directions from the navigation aid under different levels of information reliability (100%, 80% and 60%). A control condition was also used in which aids only presented a telemarketing survey and participants navigated using a map. Driver SA was assessed at the end of each experiment using a SA global assessment technique. Driver workload was collected at the same time using the NASA- TLX. Driver trust in the navigation aid information was measured using a subjective survey of initial subject trust expectations as well as a subjective rating at the close of each trial (end of

Experiment B). Across both experiments, multiple dimensions of task performance were measured.

MANOVA results for Experiment A revealed significant main effects for both ACC control mode and navigation aid type on driver performance, but no interaction effect. Findings were similar for driver SA except there was no effect of aid type. ANOVA results indicated use of the ACC system to improve driver SA and operational driver behaviors by reducing the task load in Experiment A. MANOVA results for Experiment B revealed only a significant effect of navigation aid reliability on driver performance and SA. ANOVA results revealed that perfect navigation information generally improved driving performance and driver SA for strategic driving behavior compared to unreliable navigation aid information and the control condition (task-irrelevant information). The results also revealed that drivers had higher initial trust expectations and expectation of fewer errors by the automation compared to the human. However, when participants experienced automation aid errors or inefficiency, their trust in the automation declined more sharply than trust in the human advisor. The results of this empirical work provide insight into the importance of driver SA in operational and strategic type driving tasks and associated actions. It identifies in-vehicle automation and devices as underlying factors in linkages of levels of SA to specific driving behaviors in the transactional model and serves to quantify the impact of the factors on driving performance. Validation of the proposed model and identification of other underlying factors may lead to its future use for predictive purposes.

THE EFFECT OF IN-VEHICLE AUTOMATION AND RELIABILITY ON DRIVER SITUATION AWARENESS AND TRUST

by
RUIQI MA

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the requirement for the Degree of
Doctor of Philosophy

INDUSTRIAL ENGINEERING

Raleigh

2005

APPROVED BY:

___________________________ ___________________________

___________________________ ___________________________

Chair of Advisory Committee

Dedication

To my parents, my mother Huiqiong Gui, my father Zhaoshu Ma, who gave me love and encouragement throughout my life; to my two younger brothers, Junqi Ma and Zhihua Ma, who grew up with me and shared my happiness and unhappiness; and to my wife, Na Huang, who has constantly supported me - we are looking forward to a lifetime of happiness.

.

Biography

Ruiqi Ma was born in Sichuan province in China. He was raised in Xuanhan city, Sichuan, where he completed his elementary and secondary education, graduating from Xuanhan First High School. Mr. Ma studied Industrial Automation Engineering at Beijing Science and Technology University (BSTU). He was very active in various student organizations and published many poems on the campus. Following graduation from BSTU with B.S. degree in 1993, Mr. Ma worked as an assistant electrical engineer at CERIS for four years and an electrical engineer at ABB China for two years.

In the spring of 2001, Mr. Ma began graduate study, as a research assistant, in the Department of Industrial Engineering at North Carolina Agriculture & Technology State University (NCA&T) in Greensboro, North Carolina. Mr. Ma transferred to North Carolina State University (NCSU) with a research assistant position in Raleigh, North Carolina in the fall of 2001. There he continued to study Industrial Engineering (Ergonomics). He received his Masters degree from NCSU in Industrial Engineering in 2002. Mr. Ma has served the Chinese Student and Scholars Friendship Association (CSSFA) at NCSU and was elected-President for the 2004-2005 academic year. He is currently a research assistant working with Dr. David Kaber. During his Ph.D. studies, Mr. Ma published several journal articles and gave three professional presentations.

Acknowledgement

I wish to thank a number of people for their assistance and support in the course of this work. First I would like to thank my advisor and committee chair, Dr. David Kaber. Dr. Kaber provided me with the opportunity and support to eventually complete this study. Furthermore, Dr. Kaber provided me with guidance in all aspects of my study and research at NCSU for a higher level of academia achievement. During the past four and half years at NCSU, Dr. Kaber has presented me with great enthusiasm and vision for my professional career development. Second, I would like to thank all of my other committee members, Drs. Gary Mirka, Yuan-Shin Lee and Marcia Gumpertz, as well as the Graduate School representative, Dr. Robert Buche. I appreciate your time and your suggestions to improve the quality of this work.

I would like to thank Xuezhong Wang for his assistance in building all the navigation models for my dissertation experiment. I thank Sang-Hwan Kim, who assisted me with the experimental data collection during the dissertation experiment. Also, I would like to thank Mohamed Sheik-Nainar, Jason Daly and Yongfu Ren for their assistance in the driving simulation development for my pilot study. I thank Chika Nwankwo and Carlene Perry, who assisted me with the experimental data collection during my pilot study.

Beyond this, I would like to thank the Department of Industrial Engineering at NCSU for support for my research. Lastly, I would like to thank all of the volunteers, who participated in my experiments.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACC	Adaptive Cruise Control
ANOVA	Analysis of Variance
GDTA	Goal Directed Task Analysis
HCI	Human-Computer Interaction
IP	Information Processing
LKS	Lane Keeping Systems
MANOVA	Multivariate Analysis of Variance
OOTL	Out-Of-The-Loop
RMSE	Root Mean Square Error
SA	Situation Awareness
VAS	Visual Analog Scale
VR	Virtual Reality

1 Introduction

1.1 Automation

The advancement of computer technology has led to an increased use of automation across various contexts, including power generation and process control systems, contemporary medical systems, and transportation systems. Researchers have used a number of terms to refer to automation, such as intelligent agents, expert systems and decision-support systems. Although these three terms address different aspects of automation, in this research automation is treated as a general concept encompassing intelligent agents, etc. The original use of the term automation implied automatic control in a manufacturing system, which can be traced back to 1952 (see Sheridan, 2002, page 9). According to Parsons (1985), automation can be thought of as the process of allocating activities to a machine or system to perform. More recently the term has been defined to include any use of electronic or mechanical devices to replace human labor (Parasuraman, Sheridan and Wickens, 2000). However, humans may still be involved in the process of automation.

The study of automation from a human factors perspective has recently focused on human-automation interaction in complex systems control (Sheridan, 2002). Automated systems have become so complicated that human operators may fail to effectively respond to system errors during automation failures because of a lack of understanding of manual performance of system functions (i.e., adequate skill development), or a lack of automation mode awareness. In general, the growth in complexity of automation has led to a corresponding increase in removal of operators from system control loops with the objectives of reducing errors, increasing system reliability and reducing human resources (Wickens and Hollands, 2000). Furthermore, automation has increased the amount of information that operators must process on system states and modes of automation, and a lack of appropriate system interface designs has dramatically reduced the transparency or "visibility" of automation functions to operators. An increased "distance" of human operators from direct system control (see Ma, Kaber and Chow, 2004) and decreased function visibility may deter from the successful control and diagnosis of system states during automation operation and system failure. Consequently, increases in the use of automation have often failed to yield comparable improvements in system performance (Wickens and Hollands, 2000).

A good example application of automation in our daily lives can be found in driving and contemporary in-vehicle automation. In recent years the automotive industry has designed and developed different driver assistance systems. Among these are adaptive cruise control (ACC), vision enhancement systems, lane-keeping systems (LKS), collision avoidance systems, route navigation systems, and so on (Stevens, 1997). ACC automates the driving task by maintaining vehicle speed depending upon a defined minimum vehicle following distance, by monitoring traffic (with a radar system) and instigating braking or acceleration when the detected vehicle headway distance becomes smaller or larger than the set criterion distance (Maurel and Donikian, 2001; Stevens, 1997). LKS monitor lateral vehicle position within a lane and instigate corrective steering to control vehicle position in the center of the lane (Stevens, 1997). Wickens and Hollands (2000, page 539) proposed three goals of automation, each of which serve a different purpose. They say automation is used to: (1) perform tasks humans cannot perform at all; (2) perform tasks human cannot perform very well or only at the cost of high workload; or (3) assist humans by performing tasks in which humans show limitations. In-vehicle automation, like ACC and LKS, focuses on the second and third goals identified by Wickens and Hollands. Stanton and Marsden (1996) contended that in-vehicle automation, in general, improves driver “well being”, and enhances road safety. They said that automation reduces driver stress and workload and may offer different solutions to driver errors.

1.2 Impact of automation on workload and situation awareness

Automation may have the potential to increase human operator safety and efficiency in controlling complex systems, but it may also change the operators role in controlling a system/work environment and have implications on workload and cognition. Edwards (1976) pointed out that automation might not necessarily reduce workload because it may require people to deal with additional information. Weiner (1988) said that automation might decrease workload when task responsibilities are low and increase workload when task responsibilities are high. If many automated systems are included in vehicles without consideration of human information processing capabilities, driver performance may be hindered by increased demands on attentional resources resulting from the additional task of collecting information on multiple automated system states. This situation may be further complicated by the extent of operator’ knowledge about the systems being used (Young and Stanton, 1997). For novice operators, the greater the number of information sources and tasks introduced in driving by advanced automation technologies, the greater the extent to which operator attention will be divided and workload and situation awareness (SA) may be negatively affected.

Endsley (1995a) defined SA as "the 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". SA is considered to provide a basis for decision-making and performance. With the evolution of automation, many complex, dynamic systems, which operate in uncertain environments, have been created that require the abilities of human operators to act as effective, reliable and timely decision makers. Operator SA is considered to be a crucial construct driving decision-making and performance in such environments/systems (Endsley, 1995a).

Automation can be seen to directly impact SA through three major mechanisms: (1) operator assumption of a passive role instead of an active role in controlling a system; (2) changes in complacency and vigilance associated with monitoring; and (3) changes in the quality or form of feedback provided to the human operator by the system (Endsley and Kiris, 1995). Studies of the impact of automation on SA are mixed. Wiener (1993) pointed out that the use of automation in the aircraft cockpit might improve SA by reducing system display clutter and complexity associated with manual task performance, and through the development of improved integrated display technologies. Related to this, Billings (1991) suggested that automation may also improve SA by reducing excessive workload. However, empirical research has demonstrated that high level (static) automation, such as the use of expert systems for decision making tasks, can degrade operator SA as a result of removing operators from key information processing functions (Endsley and Kiris, 1995). Endsley's (1993) research also demonstrated a degree of independence between SA and workload. Workload may negatively impact SA at very high levels. Low levels of workload may also be accompanied by low levels of SA. If workload is reduced through automation, it may not necessarily translate into higher SA.

Along these same lines, it is possible that the advanced automated technologies within automobiles (e.g., ACC and navigation assistance) may have deleterious effects on driving task SA and performance. This may be attributed to vehicle automation changing the nature of demands and responsibilities on the operator, often in ways that were unintended or unanticipated (Sheridan, 2002). The potential deleterious effects may be due to out-of-the-loop (OOTL) driver performance (or drivers being placed in a passive control role) leading to decreases in SA, and adverse influences of behavior adaptation on system effectiveness (Ward, 2000). For example, humans may exhibit over-trust in automation in automobile automation, leading to complacency, vigilance decrements and loss of SA (Sheridan, 2002).

Empirical studies have revealed the influence of in-vehicle automation on driver performance, workload and attention allocation (e.g., Parker, Rudin-Brown, and Malisia, 2003; Rudin-Brown, Parker, and Malisia, 2003). For example, results demonstrate that ACC achieves the primary goal of reducing the frequency of tailgating and the severity of rear-end collisions, and there is a significantly lower workload when the ACC is set to a long headway distance (i.e., the spacing between vehicles in time is substantial (2-3s or more)) compared to driving without ACC (Parker et al., 2003). However, it has been found that drivers may direct their attention away from the driving task when using ACC, creating an unsafe situation. The study by Rudin-Brown et al. (2003) showed a significant improvement in secondary, in-vehicle search-task performance with ACC using a long headway condition compared to driving without ACC. Drivers performed significantly fewer safe braking events with ACC using short headway (39.6% of events) and long headway (45.8%) settings, as compared to driving without ACC (63.5%). These results demonstrate that the use of ACC may primarily benefit driver workload (associated with other, non-driving related tasks) as well as performance on in-vehicle tasks other than driving. The use of ACC may also lead to unexpected increases in accidents due to driver distraction, when performing in-vehicle secondary tasks. This might counter benefits in terms of reduced tailgating and less severe rear-end collisions.

With respect to SA, a recent study by Ward (2000) indicated that in-vehicle automation, specifically ACC, appeared to achieve its goal of reducing unsafe headway distance in driving; however, there was some evidence of secondary effects of reduced SA, inferred from the observation of poorer attention to lane positioning, failure to yield to traffic, as well as slower response times to unexpected events. Ward (2000) used a performance-based measure of SA. Additional studies, like this, are needed to accurately describe the effects of in-vehicle automation on driver SA, and objective and diagnostic measures of SA need to be used.

In-vehicle devices, like cell phones, are being used more and more during driving tasks. The use of such devices represents a secondary task and may generate deleterious effects on driver SA, driving task performance and, consequently, accidents (Hancock, Simmons, Hashemi, Howarth and Ranney, 1999). For example, cellular phone usage while driving may subtract from driver visual and verbal attentional resources (e.g., watching the roadway, reading signs) degrading SA for effectively negotiating traffic, navigating, etc. With previous research in mind, it is important to know the exact SA, workload and performance effects of the introduction of automation and in-vehicle device use into driving.

1.3 Concept of situation awareness in driving

Many operational definitions of SA have been developed in the aviation domain. Aviation systems often integrate advanced automation, posing high mental demands on human operators (Billings, 1997). Although there are some similarities between the domains of flying and driving, the concept of SA has not been well defined in the domain of driving. For example, both tasks require real-time reasoning in dynamic, uncertain environments, pilots and drivers face possible information overload, and extracting relevant information from available sensors (as a basis for decision making and action) is challenging (Sukthankar, 1997). Driving, like flying, can be thought of as a dynamic system in which the system input variables change over task time. The system input variables are primarily environmental variables. They include roadway conditions, weather conditions, vehicle conditions, and driver conditions. Based on information detected on these conditions, drivers decide on a course of action that may or may not change the state of the system. Driver actions can include slowing down, accelerating, passing a vehicle, turning, etc.

In theory, the construct of SA in dynamic systems fits very well to the domain of driving. In general, driving tasks involve five time-phased information processing functions, including perception, comprehension, and projection, as well as a decision on a course of action and carrying out such action. The perception, compression and projection functions are the basis for driver SA. This cycle may or may not result in changing the state of the system, after which a new cycle of activities begins.

Driving is considered to be a complex task with many activities, some of which may be over-practiced and difficult to explain in detail (Matthews, Bryant, Webb and Harbluk, 2001). SA offers a new perspective on driving allowing for improved explanation of driving behaviors. Matthews et al. (2001) outlined multiple elements of SA that are relevant to driving, including spatial awareness, identity awareness, temporal awareness, goal awareness and system awareness. Spatial awareness refers to an appreciation of the location of all relevant features of the environment. Identity awareness refers to the knowledge of salient items in the driving environment. Temporal awareness refers to knowledge of the changing spatial "picture" over time. Goal awareness refers to the driver's intention of navigation to the destination, and the maintenance of speed and direction. System awareness refers to relevant information on the vehicle within the driving environment, which may also be viewed as a system. Gugerty and Tirre (2000) presented a similar concept of driver SA. They said drivers must maintain navigation knowledge, local scene comprehension (knowledge of nearby traffic for maneuvering), knowledge of spatial orientation, and knowledge of their vehicle's status to maintain good SA during driving.

Both Gugerty and Tirre (2000) and Matthews et al. (2001) have attempted to determine the influence of navigation (or goal) knowledge, and vehicle status knowledge on SA. System interaction knowledge is also considered to be important in a driving environment, for example, when a car traveling at a constant speed under cruise control enters a higher speed limit area, driver awareness of their vehicle speed, the speed limit and knowledge of how to set a higher speed represents good SA. There is a need for additional empirical research to provide evidence of the relative role of each form of driving knowledge on SA. Given the rapid pace of development of in-vehicle automation, study of system interaction knowledge and SA is of particular importance.

In summary, the various types of driving knowledge identified by prior research as being critical to SA, include navigation knowledge, environment and system interaction knowledge, spatial orientation knowledge, and vehicle status knowledge. These forms of knowledge can be integrated in a driver information-processing model toward an operational definition of SA in driving (see Figure 1). With a navigational goal in mind, drivers observe the driving environment. They attempt to develop the various forms of knowledge related to the driving tasks and the environment, including navigation knowledge, environment and interaction knowledge, spatial orientation knowledge, and vehicle status knowledge. This knowledge is stored in human memory (short- and long- term), as a basis for real-time decision-making. The knowledge is integrated through working memory to form driver SA. Consequently, drivers make a driving decision and implement the appropriate motor responses. Subsequent to receiving feedback on performance, drivers may form a new internal situation model and adjust their decision-making and actions leading to a new driving cycle.

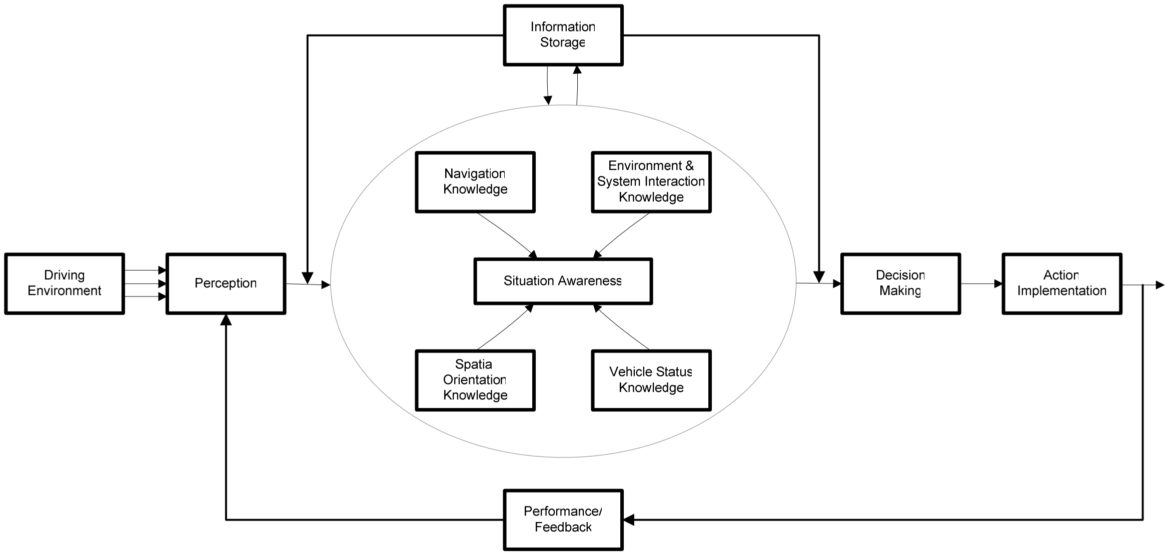


Figure 1: SA in driver information processing

As another approach to developing a model of driver behavior and SA in driving, both Ward (2000) and Matthews et al. (2001) related the three levels of cognitive functioning identified by Rasmussen (1983), skill, rule and knowledge-based behavior, to driving tasks, including operational, tactical and strategic tasks. They also related these three types of driving tasks to the three levels of SA defined by Endsley (1995a). Endsley (1995a) said the first level of SA (Level 1 SA) is based on perception of the environment. The second level of SA (Level 2 SA) is based on comprehension of the meaning of elements in the environment in relation to task goals, and the third level of SA (Level 3 SA) concerns projection of the status of elements in the near future. She said that operator achievement of higher levels of SA is dependent upon the extent to which one accurately and completely perceives states of the task environment. For example, in the context of driving, projection of the behavior of other drivers on the roadway is dependent upon accurate perception of indicators of driver intent (e.g., turn signals, brake lights, and lane changing). In operational driving tasks, drivers are engaged in actions upon vehicle actuators in order to maintain stable control. Such tasks require Level 1 SA on semi-automatic processes to ensure that the driving tasks are performed appropriately. Level 2 SA may be involved if the automatic processes “generate error messages” (leading to the need for rule-based behavior). In tactical driving tasks, there is a high requirement for Level 1 and 2 SA to facilitate local maneuvering of the vehicle in traffic streams, detecting appropriate environmental cues, and comprehending the driving situation. Tactical tasks also require short-span projection of the driving environment, probably less than the extensive projection required for strategic driving tasks (Level 3 SA). In strategic driving tasks, when navigational plans are formulated, there is a high requirement for Level 3 SA. At the time of execution, the strategic plan involves elements of Level 2 SA, in terms of perceptual integration and comprehension. There is also a small contribution from Level 1 SA, since Level 1 SA is the basis of the other two levels of SA (Endsley, 1995a; also see Matthews, et al., 2001, on page 28). These relationships between the levels of SA and the types of driving tasks are presented in a transactional model of SA in driving in Figure 2. The solid lines represent a critical link, and the dashed lines represent a potential link, between SA and driving task types in the graph, based on the literature. There has been no similar transactional model like this presented in previous studies. All of the linkages in the transactional model are based on hypotheses or inferences of other studies (e.g., Ward, 2000; Matthews et al., 2001; and Endsley, 1995a). The hypotheses were established based on the general theory of SA and the nature of the three levels of driving behavior (or driving task types). At this point in time, there is little empirical evidence to support the linkages in the new model.

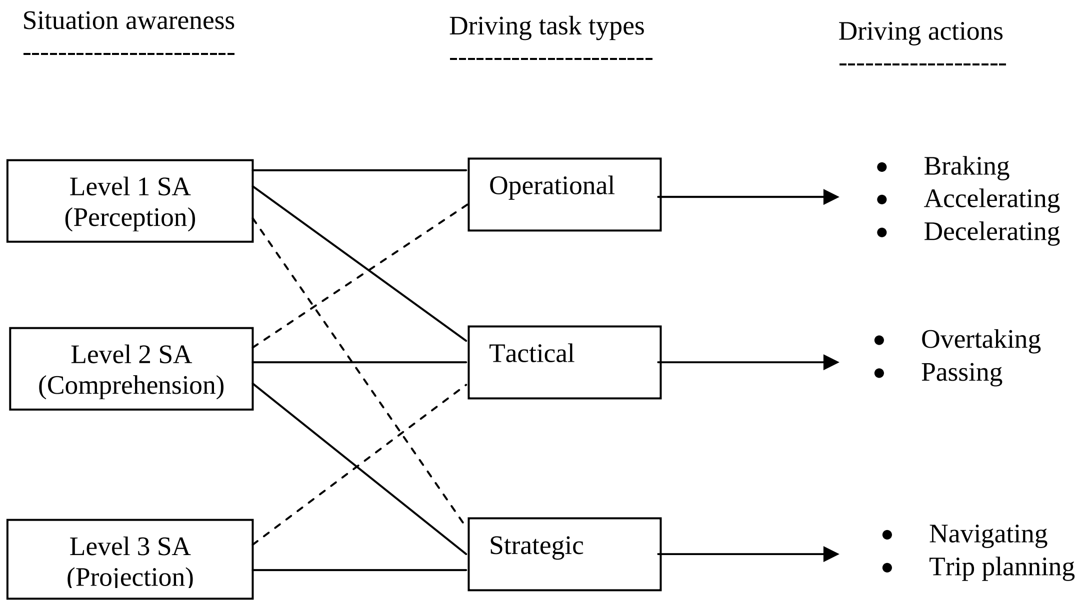


Figure 2: A transactional model of situation awareness and its potential influences on task types and driver actions

There is currently a need to develop thorough operational definitions of SA in the context of driving and to quantitatively describe the role of each level of SA in performance of the various types of driving tasks, as influenced by in-vehicle automation and devices. The model of SA in driving presented in Figure 2 allows for the specific levels of SA to be related to specific driving actions and performance. The model can serve as a basis for generating novel hypotheses and guiding the study of the role of SA in driving. Though the present research is not intended to investigate all aspects of this model, in general, it is to provide insight into how in-vehicle automation and device use may operate as underlying factors in driver SA and how changes in SA are related to specific driver actions/performance. For example, different types of automated navigation aids may influence driver Level 3 SA and, consequently, strategic driving task performance. The use of ACC may impact driver attention allocation strategies and Level

1 SA. This may lead to changes in operational task performance and specific vehicle control actions.

1.4 Secondary task performance and driving

As previously mentioned, secondary task performance, such as in-vehicle device use, during driving can have a negative effect on driver SA and performance. Jerome, Ganey, Mouloua and Hancock (2002) said that one of the central concerns for driving today is the effect of in-vehicle devices on driver performance and safety. Here, the terminology "in-vehicle devices" is used to refer to any device a driver can manipulate while driving, which is not directly related to the driving task, for example, the car radio or a cellular phone. The use of cell phone and wireless communication devices has increased at an exponential rate over the past two decades (Edwards, 2001). With more and more cell phone usage during driving, it is critical to know if cell phone conversations in cars increase driver workload, and decrease SA, ultimately leading to decreases in task performance.

A number of researchers have observed that it is not easy to quantify the extent to which driving performance is compromised when a secondary task, such as using a cell phone, is taking place at the same time. This is due to differences in driver' abilities and skills, differences in driving conditions, and various levels of complexity of in-vehicle devices and tasks (e.g., dialing a cell phone, answering calls, talking). In general, one would expect a cell phone conversation, when driving, to cause the same disruption for a driver (in terms of achieving and maintaining SA), as having a conversation with a passenger. Both activities compete for limited driver mental resources (visual and verbal attention and processing (see Navon and Gopher, 1979). However, the cell phone conversation may be worse in terms of impacting driver SA and performance since the caller cannot visualize the driving situation and assist the driver, and the driver may need to use one hand to hold the phone (a hand-held phone) while driving. A study by Chen and Lin (2003) compared driving situations with and without a secondary conversation using a driving simulator. Results indicated that the use of a mobile phone while driving might have adverse implications for driving safety. The study showed that drivers adopted several approaches to reducing the cognitive workload caused by the phone conversation. Participants compensated for a need for increased reaction time by increasing headway distance to other cars and decreasing driving speed during the dual-task situation (driving and talking). Chen and Lin also observed an increase in missed brake responses, which seemed to be caused by a loss of attention to the roadway in the dual-task situation, which could have led to accidents. Furthermore, the dual-task driving test indicated a loss of attention in perceiving information/warnings presented on road signs.

Another study by Gugerty, Rando, Rakauskas, Brooks, and Olson (2003) assessed differences between remote and in-person communication during driving. Results indicated that the pace of the in-person and remote verbal interactions significantly differed. Drivers talking with remote partners generated longer pauses than drivers talking with in-person partners, suggesting that drivers engaged in remote verbal interactions were modulating their conversation in order to maintain adequate driving performance. SA was also assessed in this study using location-recall probes requiring participants to indicate the locations of surrounding traffic. Performance probes were used and required participants to avoid nearby hazardous vehicles. Scene-interpretation probes required participants to identify cars that were driving dangerously. Gugerty et al. found that SA was significantly degraded when performing the driving task while talking with a partner as compared to only driving a car. However, the amount of degradation in SA during in-person and remote interaction did not differ significantly.

Several studies (Hancock, Simmons, Hashemi, Howarth and Ranney, 1999; Hancock et al. 2002) have indicated that there is a slower response to traffic light changes in the presence of an in-vehicle distracter (cell phone conversation), and have affirmed deleterious effects of competing tasks on performance at crucial points in a driving maneuver using basic automatic-transmission cars. This research also suggested a possible decrement of SA (Level 1 SA, perception, Level 2 SA, comprehension, and Level 3 SA, projection), as a result of cell phone usage during driving.

In general, it appears that secondary (distracter) tasks, such as cellular phone conversations, during driving may compete for limited driver mental resources, causing possible SA decrements and decreases in human performance. However, there remains a need to make direct, objective assessment of cell phone use while driving on the various levels of driver SA, including perception, comprehension and projection in normal driving circumstances.

1.5 Trust in automation

1.5.1 Trust and technology

The concept of trust originates from human-human interaction. It refers to the expectation of, or confidence in, the behavior of another. Trust is based on the probability that one party attaches to cooperative or favorable behavior by other parties (Hwang and Burgers, 1997). Although trust has long been identified as a critical factor in many non-technical fields, researchers are just beginning to study it in the context of human interaction with technology. Streams of research on trust can be found in the fields of sociology, psychology, marketing, ergonomics, human-computer interaction (HCI), and electronic commerce (Corriotore, Kracher and Wiedenbeck, 2003). Although trust has been studied in a variety of disciplines, each of these disciplines has produced its own concepts, definitions and findings. In fact, even within a given field, there is often a lack of agreement and focus of research effort on trust (Lewicki and Bunker, 1995). The outcome is a multi-dimensional family of trust concepts, each with a unique focus.

Despite the eclectic nature of trust research, researchers from every discipline do acknowledge the value of trust. Trust enables people to live in risky and uncertain situations (Mayer, Davis and Schoorman, 1995). It provides the means to decrease complexity in a complex world by reducing the number of options one has to consider in a given situation (Lewis and Weigert, 1985). Trust would not be needed in human-human or human-automation interaction if we could undertake actions with complete certainty and no risk.

1.5.2 Defining trust in automation and factors

The degree of trust a human places in a machine or automation is one of most critical factors that influences complex system operator use of decision support systems. The introduction of automation into complex systems such as transportation systems, process control systems, medical systems, and robotic systems has led to a redistribution of operational responsibility between human operators and computerized automated systems. Research in ergonomics has examined how human trust has been established, maintained, lost and regained in the development of human-machine systems, in which operational control is shared by the human and the automation. Muir (1994) characterized trust as an intervening variable that mediates user behavior with computers. Muir and Moray (1996) argued that trust in automated machines is based mostly on user perceptions of the expertise of the machine or automation in properly performing a function that may have been previously performed by the user. This trust influences the resulting behavior of operators and overall task performance. Furthermore, according to HCI studies, users who have low knowledge or self-confidence in a task situation tend to trust a computer system because it provides expertise that the user lacks (Lee and Moray, 1992; Kantowitz, Hanowski and Kantowitz, 1997). Conversely, when users are familiar with, and self-confident in, a task situation they have a higher standard for acceptance of advice from automation and, therefore, a higher threshold for trust (Kantowitz et al., 1997). Users have also been shown to trust a computer if they have tried and failed to solve a problem on their own (Waern and Ramberg, 1996). Related to this research, Sheridan (2002) makes a distinction between different meanings of the term ‘trust’ in the context of human-automation interaction. Specifically, he distinguishes between trust as an effect or outcome of certain automation characteristics (e.g., reliability) and trust as a cause of operator’ behavior when utilizing automation. That is, human operator trust in automation, based on system reliability, significantly affects whether and how automation is used.

Recently, automation researchers have begun to make a distinction between automation trust and automation reliance (Wickens and Hollands, 2000). According to Wickens and Hollands, automation trust is defined in terms of subjective measures, such as user confidence ratings in the automation or their verbal estimates of automation reliability. In contrast, automation reliance is defined in terms of performance or behavioral measures such as automation utilization and efficiency. These definitions of automation trust and reliance are adopted in this study. Clearly, such a distinction has important implications for assessing and understanding the impact that different levels of automation reliability may have on operator trust in automated aids, as well as for the design of automation interfaces aimed at improving trust calibration and subsequent system performance (Wiegmann, Rich and Zhang, 2001).

1.5.3 Automation reliability and operator trust

Empirical studies of trust in automated machines show that performance and trust increase following a similar curve, as long as there are no automation errors (Lee and Moray, 1992). Failure or errors in automated systems may arise from control algorithms that are not optimal for all operating circumstances or communication/interaction between the human operator and automation. The sophistication of control algorithms depends on engineering technology and analytical tools, as well as modeling of human decision processes. With respect to the second issue, human operators may or may not respond to automation in the way that system designers expect. Researchers have found human operators may underutilize and over rely on automation depending upon its capabilities and reliability (Dzindolet, Peterson, Pomranky, Pierce and Beck, 2003; Parasuraman and Riley, 1997).

Machine/automation errors can have a strong, degrading affect on operator trust. The magnitude of an error is also an important factor in loss of trust (Muir and Moray, 1996; Kantowitz, Hanowski and Kantowitz, 1997). Lee and Moray (1992) found that errors from automation led to a sharp drop in trust roughly proportional to the magnitude of the error. If the error was not repeated, performance recovered immediately, but recovery of trust to prior levels occurred over a longer time. An accumulation of small errors also decreases trust (Lee and Moray, 1992; Muir and Moray, 1996) and these small errors appear to have a more severe and long-lasting impact on trust than a single large error. Even in the face of automation errors, a user may continue to trust a computer system in certain situations, for example, if the errors are predictable (Muir and Moray, 1996). If the user is able to understand and compensate for the errors, recovery of trust can occur even when small errors continue (Lee and Moray, 1992; Muir and Moray, 1996). Errors encountered in one function of an automated system can lead to distrust of related functions, but do not necessarily generalize to an entire system (Muir and Moray, 1996).

Related to automation error/inefficiency and trust, most studies have examined how trust develops when interacting with automation of a single reliability level. Results of the few studies that have systematically varied automation reliability levels are mixed. Dzindolet, Pierce, Beck, Dawe and Anderson (2001) required participants to view slides of battlefield terrain and to indicate the presence or absence of a camouflaged soldier through the assistance of an automated decision aid. Their results suggested that operators were insensitive to differences in automation reliability. Wiegman et al. (2001) examined the effects that different levels of, and changes in, automation reliability have on users' trust of automated diagnostic aids. Both subjective measures (perceived reliability of the aid) and objective measures of performance (concurrence with the aid's diagnosis and decision, and time of automation reliance) indicated that users were sensitive to different levels of aid reliabilities.

1.5.4 Trust in human aids versus automation

Another interesting line of trust research focuses on differences in trust in complex system controllers when relying on information from an automated aid (e.g., computer) versus a human aid. Human beings have personal experiences that lead to calibration of trust in other people. However, humans must develop mental models of automation in task performance through limited training experiences, and these mental models lead to trust calibration for interacting with the automation. A study by Dzindolet et al. (2001), in which participants were required to view slides of battlefield terrain and indicate the presence or absence of a camouflaged soldier through the assistance of an automated or human decision aid, revealed a significant bias toward automation in terms of complex system operator trust. The automated aid was perceived as more reliable than a human aid. Trust in the automated aid appeared to vary depending upon the level of reliability and was related to the functions of the aid. However, according to Dzindolet, Pierce, Beck and Dawe (2002), higher perceived source reliability or credibility of the automaton aid was not reflected in objective automation (use) reliance strategies, as participants in their experiments showed a strong tendency toward self-reliance. In summary, humans may trust machines more than other humans for aiding information processing; however, humans exhibit greater self-reliance than machine use when machines are unreliable.

Research has also revealed that people are more sensitive to the errors made by automation than human aids, leading to a very rapid weakening of credibility and a swift decline in trust of automation (e.g., Dzindolet et al., 2001; Wiegmann et al., 2001). There is a need to know how varying automation reliability influences operator trust in automation, and if there is a difference in trust of an unreliable human-aid versus automation-aid in realistic tasks. There is also a need to explain any differences, or to identify underlying factors. When levels of automation reliability vary, this may pose a different mental demand on human operators as a result of having to monitor both system states and automaton states. Furthermore, it may influence operator SA. Under varying reliability automation, lower reliability conditions require more mental attention, reducing operator perception, comprehension and projection of system states and environment knowledge (i.e., SA). No studies have investigated the impact of automation reliability on direct, objective measures of SA.

2 Problem Statement

In general, knowledge of SA in driving under various scenarios involving in-vehicle automation or device use is not complete. Although the use of telematics and secondary-tasks appear to have the potential to significantly degrade driver performance, there have been few empirical studies of the impact on driver SA as a potential cause of performance problems. Some studies reviewed here measured aspects of SA; however, they may not accurately reflect changes in driver perception, comprehension and projection, because of the use of inferred measures (e.g., Ward, 2000). SA has been considered to a limited extent in the analysis of driver behavior, but little work has empirically examined the cognitive construct when drivers are using advanced automated technologies (e.g., ACC) or personal communication devices (e.g., cellular phones). Driver SA and performance may be hindered by an increased processing load resulting from the additional tasks of collecting information about automated system states and concentrating on cell phone conversations. There has been no study of SA when drivers are using advance-automated technology and cell phones in combination. There is a need to make direct, objective assessment of SA under these circumstances.

There are few studies that have systematically varied automation reliability levels in investigating human trust in automated systems, such as in-vehicle navigation technologies for guidance. There is a need to understand how varying levels of automation reliability influence operator trust in automation and SA, and if there is a difference in trust in human aids versus automation aids in realistic tasks. There is also a need to explain why any differences may exist. No studies have considered the impact of in-vehicle automation reliability on SA. If drivers perceive different reliabilities of in-vehicle automation systems, they may allocate more attentional resources from an already limited source in order to monitor automation states. Therefore, there may be a negative influence on driver SA.

The objectives of the current research were to: (1) investigate the implications of ACC and cellular phone use in driving on a direct and objective measure of SA; (2) investigate the effect of varying reliability of in-vehicle automation (navigation aids) on driver SA and trust; and (3) assess differences in human trust in a human aid versus an automation aid in a simulated driving task. The study was expected to detail aspects of the proposed transactional model of levels of SA in various driving behaviors and activities, as influenced by automation and in-vehicle device use. Since driving is an over-practiced task, introducing ACC and different levels of reliability of a navigation information aid in a simulated vehicle was expected to allow for assessment of potential behavior adaptation in driving operations in order to address varying attentional loads imposed by the ACC and navigation aid. It was expected that drivers would develop different strategies in making decisions in the complex control task due to workload reductions created by the ACC and workload increases due to varying reliability of the information aid. The research was also expected to allow for observation of a possible connection between SA and accurate decisions in critical driving situations.

3 Pilot Study

A pilot study was conducted to address the first major need identified in the problem statement, specifically making assessment of the effects of ACC and cell phone use in driving on SA, and assessing the competition of multiple driving and communication tasks for limited mental resources in terms of driving performance. This required developing a driving simulation and experimental devices for use in experimentation and data collection and developing a valid operational definition (objective measure) of SA.

Contrary to Ward’s (2000) inferences on SA in driving, it was hypothesized that use of the ACC system would improve driver SA under normal driving conditions (i.e., no unexpected events or hazards). We expected the ACC to reduce task load in terms of the need to monitor for and implement vehicle speed changes and, thereby, free-up cognitive resources for perceiving the driving environment. The use of the ACC system was accordingly hypothesized to decrease driver perceived mental workload as a result of relieving them of the need for continuous speed and headway distance control. Based on Ward’s (2000) findings, the ACC system was also hypothesized to provide better task performance than no-ACC driver speed control and headway distance control because of the potential for driver boredom and vigilance decrements (not paying attention to speed limits, etc.) over extended periods of manual control/no-ACC.

The cell phone conversation during driving was expected to compete for limited driver mental resources and to increase driver perceptions of workload and, as Gugerty et al. found, to decrease SA. Based on the results of Chen and Lin (2003) and Hancock et al. (1999, 2002), the cell phone use was also expected to degrade driver task performance.

Finally, based on the findings of Rudin-Brown et al. (2003), the combined use of the ACC and the cell phone was expected to create a situation in which the driving workload relief provided by the in-vehicle automation would lead to increased driver concentration on the secondary task (the cell phone conversation). This situation was expected to degrade SA and overall driving performance. One concern with respect to this hypothesis was that prior work examining the effects of in-vehicle highway systems on driver secondary-task performance used simulations in which the secondary task occurred continuously during driving versus intermittently, like real cell phone conversations. It was suspected that intermittent cell phone calls in this study would be less distracting to drivers than continuous secondary-task performance (e.g., eating while driving) and would constitute a more conservative assessment of the SA effects of the automation and in-vehicle device use.

3.1 Methods

3.1.1 Simulation and driving task

The simulation used in this study was a medium fidelity, 3-dimensional representation of a dynamic freeway-driving environment. The terminology, “medium fidelity” is used because the simulation was presented using a virtual reality (VR) system, including a stereo display. User control inputs occurred through realistic automobile control interfaces, including a physical steering wheel, and physical gas and brake pedals (see Figure 3). However, this simulator was a fixed-base setup providing no kinesthetic motion and there was no interactive traffic in the simulation (i.e., vehicles appeared in the rear-view mirrors of the user's car, but they did not pass or cross the participants vehicle). The simulation required participants to drive a virtual car and perform a following task, which involved changes in speed and lateral position. The simulation environment included a four-lane highway presented from an egocentric viewpoint inside a driver's sports vehicle. The roadway was marked with conventional lines. There were also many types of signs along the sides of the highway, including: "pedestrian crossing", "slow", "deer crossing", "railroad" and "speed limit". The environment included buildings, grass, rivers and street lights (see Figure 4). All objects in the virtual environment were modeled to scale and presented with rich, realistic textures.

Figure 3: Experiment set up

Figure 4: Driving simulation

Participants were asked to drive on the roadway, maintain their vehicle in the right-hand lanes (of the four-lane freeway), keep their vehicle in the middle of a particular lane, and follow the lead vehicle. They were also asked to observe all road signs. Participants were exposed to ACC or no-ACC control modes, of which they were informed in advance. The ACC automated the driving task by maintaining vehicle speed depending upon a defined minimum vehicle following distance (approximately 2.4 s) and a maximum travel speed (80 mph). Certain participants were also required to talk on a cell phone with a remote partner. The cell phone call was considered to be a secondary (distracter) task, in which an experimenter asked participants a number of arithmetic questions (10 problems per call, including a single digit multiplied by a two-digit number or multiplication of two numbers with one digit, each).

Participants drove for roughly 25 minutes during each trial. The freeway was approximately 25 miles in length, including straight-aways and curves configured in a giant loop. The average speed of the lead vehicle in all trials was 60 mph.

3.1.2 Experimental design

The independent variables for the experiment comprised the ACC control modes (active or inactive) and the cell phone use (conversation or no conversation) condition. The ACC condition was manipulated within-subjects because of the driving experience of participants; therefore, we expected little or no carry-over effect across conditions. The cell phone conversation condition was manipulated between subjects in order to reduce possible condition carry-over effects, as participants might have become more proficient at responding to the arithmetic questions while driving. Each participant in the cell phone condition groups completed two trials under each ACC setting. In total, all participants completed four 25-minute sessions during the experiment.

The dependent variables for the experiment included driver SA. Participant perception (Level 1 SA), comprehension (Level 2 SA) and projection (Level 3 SA) were measured using the Situation Awareness Global Assessment Technique (SAGAT). The SAGAT is a simulation-freeze technique in which SA queries are posed to complex system operators (in this case, drivers) on the state of a simulation at various points in time (Endsley, 1995b). In the present experiment, the driving simulation was frozen at 7, 14 and 21 min. into a trial. During a freeze, the simulation display screens were shutdown and participants moved to an adjacent workstation at which they found a pencil and SA questionnaire sheet. Each questionnaire presented a sample of nine SA queries from a pool of 27 queries targeting all levels of SA. Each questionnaire included three Level 1, 2 and 3 SA queries. Participants were required to recall car locations and colors or traffic signs they had passed. They were required to identify any necessary driving behaviors (acceleration, braking and turning) to improve the accuracy of their following position behind the lead car. They also projected times to certain events, such as the time to the next turn or to pass the next sign in view, etc. (see Appendix A for an example questionnaire). A goal directed task analysis (GDTA), using the methods described by Endsley and Jones (1995), was conducted to identify major goals required for accomplishing the lead-car following task, subgoals that are essential for meeting the overall goals, major decisions that are associated with each subgoal, and SA requirements for accomplishing the task. Consequently, the SA requirements that represent a level of information processing (perception, comprehension, or projection) were used to develop SAGAT queries (see Appendix B for the GDTA conducted for the pilot study). There was no time limit on participant responding to queries. After participants completed a questionnaire, they returned to the driving simulation workstation and continued the virtual task from exactly where they left off. The SA response measures for each trial included the average percent correct participant responses to Level 1, 2 and 3 queries and a total SA score across all three questionnaires. Participant answers to the SA queries were graded based on "ground truth" on the simulation recorded by the VR system at the time of the SAGAT freezes.

Subjective workload was also subjectively measured after each session by using a mental demand rating scale with anchors of “Low” and “High”. Participants marked an “X” on the scale at the position they felt most accurately represented the demand for the trial. The response measure was the distance from the “Low” anchor to the participant’s rating divided by the total length of the scale.

Finally, task performance was measured in terms of participant accuracy in lane maintenance and tracking lane changes by the lead car, as well as tracking lead car speed and maintaining safe headway distance (the optimal range was defined as 8 to 25 meters) in the following task. Task performance was recorded automatically by the VR computer system at every second during the simulaiton trials. The root mean square error (RMSE) for the headway distance and following speed, as well as lane tracking and maintenance on the straight and curve segments of the driving loop, were caculated for each trial.

3.1.3 Apparatus

The driving simulation was programmed using Visual C++ and the Virtual Environment Software Sandbox (VESS) was used as a real-time VR engine. Participants wore stereographic goggles to view the VR in 3-D. A Motorola T720 cell phone was used for all phone conversations during the experiment. An experimenter called the cell phone during trials from a landline, using a speakerphone in an adjacent lab room, which could not be seen by participants, nor could they directly hear the voice of the experimenter.

3.1.4 Participants

Eighteen college students were recruited for the study. Half the participants were assigned to a group required to have cell phone converstions while driving. All participants were required to have 20/20, or corrected to normal vision, and at least one year of driving experience. Nine males and nine females participated in the actual experiment with an equal number assigned to each cell phone condition. The average age of the participants was 26.6 years, and there was an average of 6.11 years of driving experience.

3.1.5 Procedure

Each participant completed the entire experiment in one day according to the following procedures: (1) 20 min. of instruction on the driving simulation; (2) 20 min. of training in the simulation driving task under a manual/no-ACC control mode (without cell phone use); (3) 15 min. of instruction of the SA questionnaire and subjective workload rating scale to be administrated during experimental trials; and (4) four 25-min. trials, including the three SA questionnaires and the summary workload rating with intervening 5-min. breaks between trials. Participants were instructed to concentrate on the driving task and allocate whatever residual attention they may have to other tasks (i.e., the cell phone response). If participants were assigned to the cell phone conversation condition, calls were received at 3, 10 and 17 min. into each trial, and each call lasted slightly less than 2 minutes. Figure 5 presents the schedule of events during each experiment trial. The experiment lasted 3.5 to 4 hours for each participant.

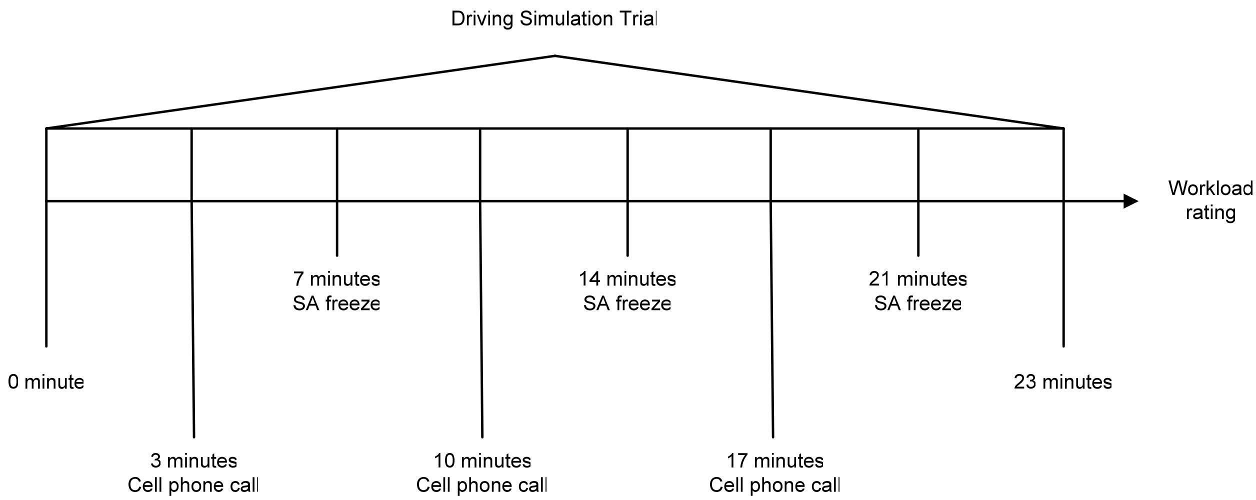


Figure 5: Schedule for secondary (distracter) tasks and SA freezes during test trials

3.2 Results and discussion

3.2.1 Driver SA

Figure 6 presents the mean Level 1, 2, 3 and total SA scores for both ACC control mode and cell phone conversation condition. The plot reveals that, on average, drivers exhibited better SA when the ACC control was active and no-cell phone conversation took place.

Analysis of Variance (ANOVA) results on driver SA indicated that, in general, the ACC control mode and cell phone conversation conditions were influential in the percentage of correct responses to SA queries during trials. There was a significant effect of ACC control mode on Level 1 SA ($F(1,16)=18.68$, $p=0.0005$) with greater perceptual knowledge of the driving environment occurring when the ACC control was active. There was no interaction effect of the ACC control mode and cell phone conversation on Level 1 SA.

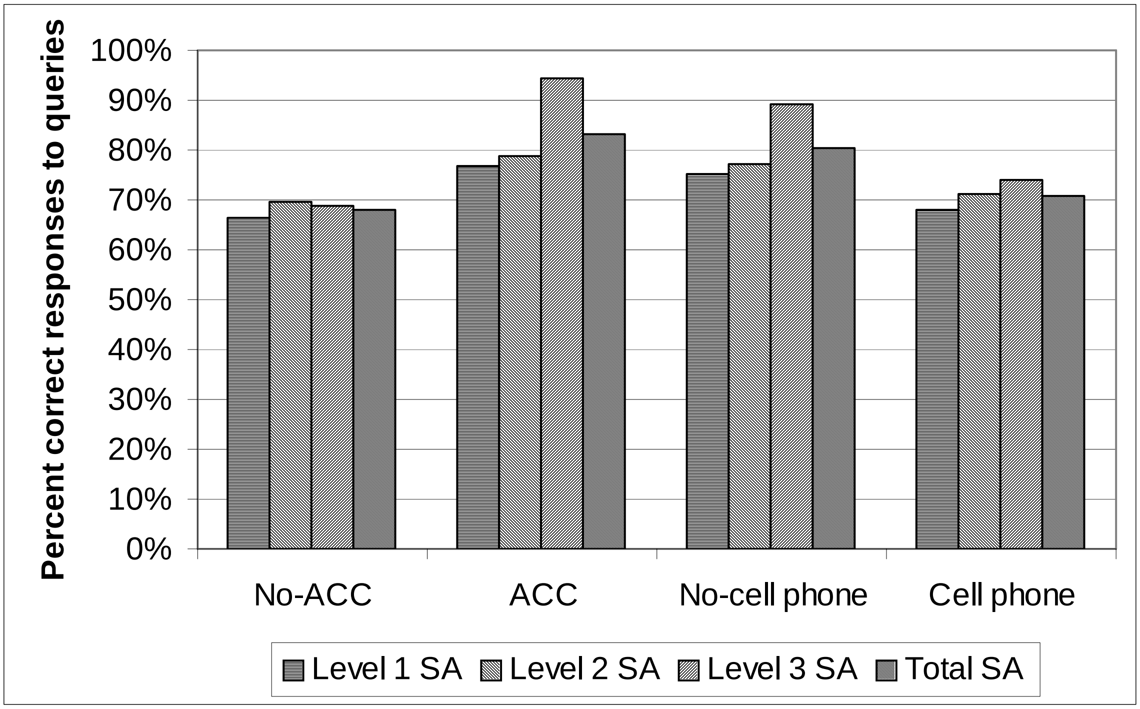


Figure 6: Mean SA for ACC control and cell phone conversation conditions

ANOVA results revealed significant Level 2 SA effects of the ACC control mode ($F(1,16)=22.22$, $p=0.0002$) and cell phone conversation ($F(1,16)=5.15$, $p=0.0375$). Drivers demonstrated significantly greater comprehension of the driving environment when using the ACC control. As we hypothesized, the cell phone conversation degraded driver SA and there were significantly higher scores for Level 2 SA when no cell phone conversation took place. There was no interaction effect of the ACC control mode and cell phone conversation on Level 2 SA.

ANOVA results revealed significant Level 3 SA effects of the ACC control mode ($F(1,16)=121.73$, $p<0.0001$) and cell phone conversation ($F(1,16)=36.26$, $p<0.0001$). Drivers demonstrated significantly greater ability to project states of the driving environment when using the ACC control. Similar to the results on Level 2 SA, there were significantly higher scores for Level 3 SA observed when no cell phone conversation took place; that is, the cell phone conversation degraded driver projection of states of the driving environment. There was also a significant interaction effect of the ACC control mode and cell phone conversation condition on Level 3 SA ($F(1,16)=15.22$, $p=0.0013$). Figure 7 presents the ACC system and cell phone conversation interaction effect on Level 3 SA. Tukey's test revealed significantly higher ($p<0.05$) projection scores when the ACC was active across both cell phone conditions than when the ACC was inactive and cell phone conversations did not take place. It also revealed significantly higher ($p<0.05$) Level 3 SA scores when the ACC was inactive and no cell phone conversation took place, as compared to no ACC with cell phone conversations.

Unlike the results on Level 1 and 2 SA, the findings presented here suggest that drivers may not be able to continue to make accurate projections of the driving situation when posed with secondary distracter tasks. Among the various stages of information processing (IP) encompassed by the construct of SA, the stages of perception and comprehension may place relatively lower demands on human mental resources, as compared to projection, and consequently drivers may be able to address such demands even when resource competition occurs (i.e., a cell phone call occurs). For system-state projection, humans may not be able to manage information on the driving environment and from a cell phone conversation, and to simultaneously make accurate judgments on the future of the driving situation.

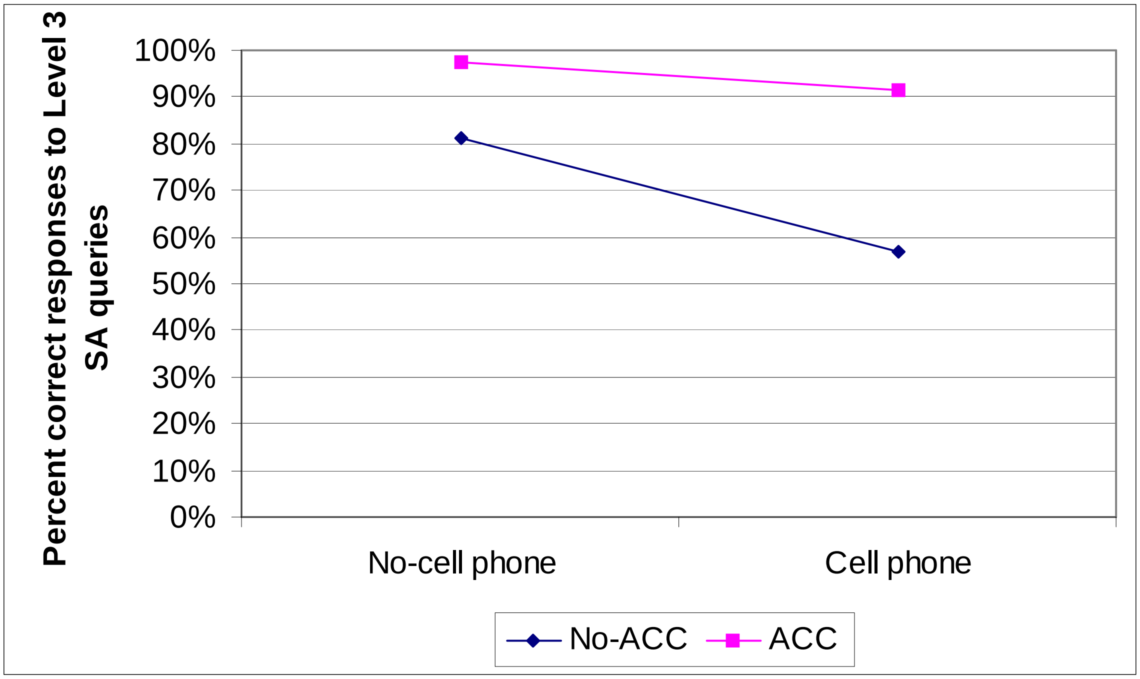


Figure 7: ACC and cell phone conversation interaction effect on Level 3 SA

ANOVA results also revealed overall SA (or the total SA score) to be significantly affected by the ACC control mode ($F(1,16)=118.38$, $p<0.0001$) and cell phone conversation condition ($F(1,16)=20.75$, $p=0.0003$). There were significantly higher scores for total SA when the ACC control was active. There were also significantly higher scores for total SA when no cell phone conversation took place during the trials. There was no interaction effect of the ACC control mode and cell phone conversation condition on total SA.

In summary, these findings support the general notion that introducing the use of automation in vehicles under normal driving conditions allows for improvements in driver SA by reducing driver task load in terms of the need to monitor for, and implement, speed changes. As expected, the results on SA also supported the contention that the cell phone conversation would degrade driver comprehension and projection of states of the driving environment, and overall SA. Although the ANOVA results did not reveal an effect of the cell phone conversation on Level 1 SA, this may due to the short duration of cell phone conversation time during the experiment (approximately 1.5 min. for each call).

3.2.2 Driving workload

ANOVA results revealed subjective ratings of mental demand in the driving task to be significantly affected by the ACC control mode ($F(1,16)=68.46$, $p<0.0001$) and cell phone conversation condition ($F(1,16)=8.54$, $p=0.01$). Figure 8 presents the mean percent mental workload for both the ACC control mode and cell phone conversation conditions. The mean percent mental demand was significantly greater when there was no ACC control. There were also significantly greater perceptions of mental workload when cell phone conversations took place (and recall this condition was manipulated as a between-subjects variable). There was no interaction effect of the ACC control mode and cell phone conversation condition on workload.

The findings on workload support the hypotheses that under normal driving conditions the use of the ACC and cell phone would decrease and increase driver mental workload, respectively. In this study, the ACC system and cell phone appeared to have comparable influences on mental workload (compare the bars in the graph in Figure 8). The findings presented here support an advantage of the introduction of in-vehicle automation during normal driving conditions, and suggest the importance of limiting cell phone usage. Related to the hypothesis on the interaction effect of the use of in-vehicle automation and cell phone calls on the demand for cognitive resources, it is possible that the automation did provide workload relief, but that the cell phone conversations exploited this, consequently, washing-out any significant effect across conditions.

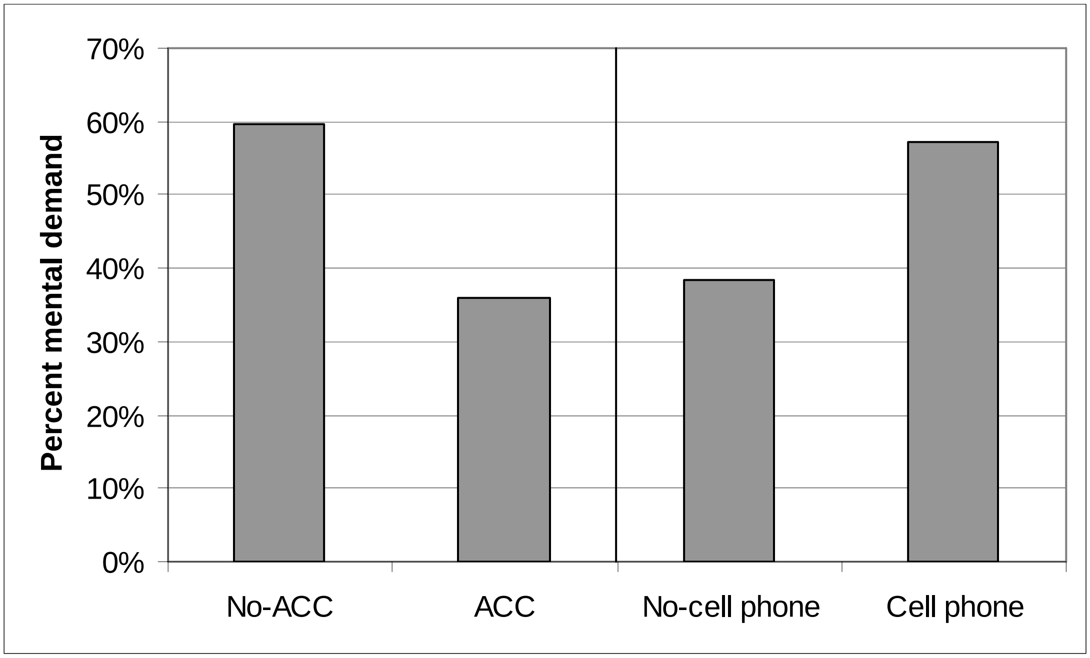


Figure 8: Mean workload rating for ACC control and cell phone conversation conditions

3.2.3 Driving performance

In general, performance results indicated that the ACC system was influential in vehicle control, but that the cell phone conversation condition was not. This observation may be attributable to the concern that the cell phone conversations were intermittent and did not pose a continuous secondary task load on drivers throughout trials. The specific findings on headway distance, speed control and lane maintenance are presented here.

3.2.3.1 Headway distance

ANOVA results revealed a significant effect of the ACC control mode on variation in headway distance ($F(1,16)=42.53$, $p<0.0001$). Figure 9 presents the RMSE of headway distance for both the ACC control mode and cell phone conversation conditions. Drivers appeared to allow significantly greater deviations in headway distance when the ACC control was inactive, possibly suggesting a perceived need for greater caution at the test speeds, or limited driver confidence in their ability to quickly react to lead vehicle speed changes. There was a trend for greater headway deviations when drivers used the cell phone but the effect of the secondary-task on driving performance did not prove to be statistically significant. There was no interaction of the ACC control mode and cell phone conversation condition in terms of headway distance.

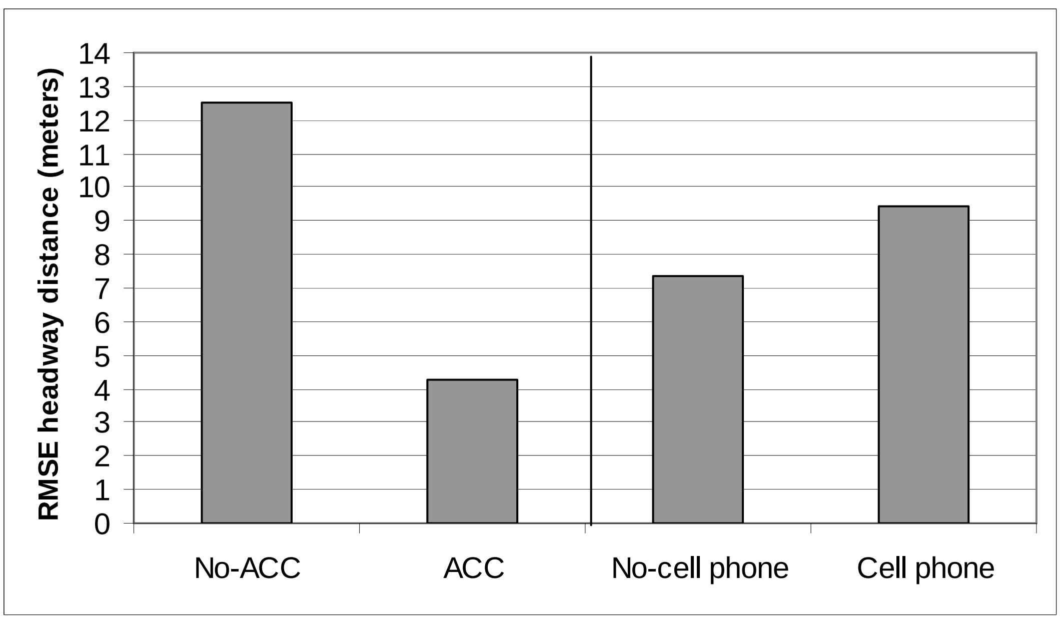


Figure 9: RMSE of headway distance for ACC control and cell phone conversation conditions

3.2.3.2 Following speed

ANOVA results revealed a significant effect of the ACC control mode on variations in driver following speed (when tracking the lead vehicle in the simulation) ($F(1,16)=111.95$, $p<0.0001$). Figure 10 presents the RMSE of following speed for both ACC control mode and cell phone conversation conditions. There were significantly greater deviations in following speed with no ACC/manual control. Again, there was a trend for worse speed control when using the cell phone, but the difference among the conditions was not significant at the selected alpha criterion of 0.05. There was no interaction of the ACC control mode and cell phone conversation condition in terms of variations in following speed.

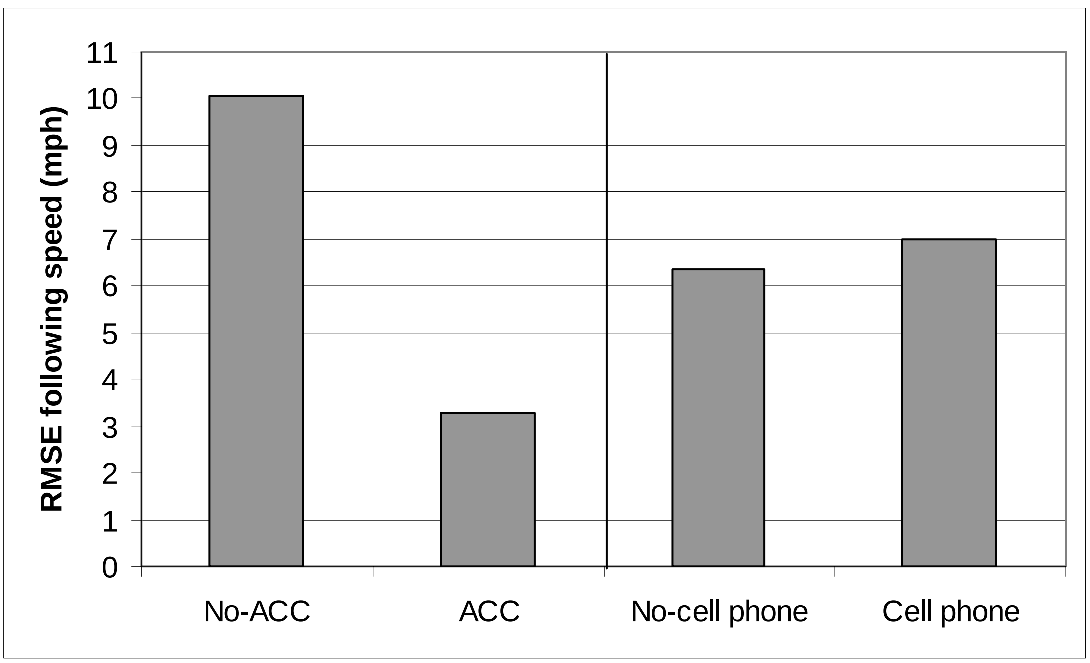


Figure 10: RMSE of following speed for ACC control and cell phone conversation conditions

3.2.3.3 Lane maintenance on curves

ANOVA results on driving performance indicated that the ACC control mode had a marginally significant effect ($F(1,16)=4.36$, $p=0.0526$) on driver lane maintenance only when negotiating curves as part of the simulated freeway. There was no significant effect of ACC control on lane maintenance on straight-aways. There was a trend for greater lane maintenance deviations on curves when the ACC control was inactive. There was no main effect of cell phone use on lane maintenance or an interaction of the ACC control mode and cell phone conversation condition in terms of straight-away or curved lane maintenance deviations.

In summary, it was hypothesized that the ACC system would facilitate better driving performance, including speed and headway distance control, and lane maintenance. The findings presented here generally support the use of ACC control to improve driving performance under normal driving conditions. The results did not support the hypothesis, based on Chen and Lin (2003), that cell phone conservations during driving would decrease task performance. However, once again, this may be attributable to the short duration of the cell phone conversations during trials, as part of the present study. Although there were three cell phone conversations during a single test, they were brief and the total cell phone conversation time for any trial was much shorter than the total driving time (approximately 5 min. versus 25 min.). Therefore, the cell phone condition did not pose a consistent secondary-task demand on drivers, potentially subtracting from performance.

3.2.4 Correlation analyses

Simple correlation analyses were conducted in order to identify any significant relationships among SA, workload and secondary-task performance (percentage of correct responses to arithmetic problems during cell phone conversations), as well as driving performance. A Pearson correlation coefficient revealed a significant negative linear association between the total SA score and subjective workload ratings ($r = -0.716$, $p<0.0001$). There were also highly significant negative linear associations between workload ratings and percent correct responses to queries on each level of SA. These additional findings add strong support to our contention that the in-vehicle automation off-loaded drivers in terms of monitoring motor-control tasks and allowed for greater perception, comprehension and projection of driving environment states (i.e., SA improved as workload decreased through the use of automation).

A Pearson correlation coefficient revealed a significant negative linear association between workload ratings and secondary-task performance ($r = -0.447$, $p=0.0063$) (i.e., a positive correlation among subjective and objective workload measures). As ratings of mental demand in the driving task increased, secondary-task performance decreased, and vice versa. This finding further demonstrates the mental resource competition among driving tasks and cell phone use. It also supports the use of secondary-task measures of mental workload in driving simulations.

Finally, Pearson correlation coefficients revealed significant negative linear associations between total SA score and variations in headway distance ($r = -0.49882$, $p=0.002$) and following speed ($r = -0.5498$, $p=0.0005$). There was also a significant negative linear association between Level 3 SA and variations in headway distance (r = -

0.66129, $p<0.0001$) and following speed ($r = -0.70919$, $p<0.0001$). All these findings indicate positive associations of the construct of SA and driving performance; that is, as SA increased, the RMSE in headway (from the optimum range) and speed decreased. They support the linkages among all levels of SA and operational driving task actions, and in particular links among driver projection of environment states and operational behaviors, in the transactional model of SA in driving (Figure 2). This correlation evidence can also be considered validation of our operational definition of SA (objective measure) based on the GDTA of the driving simulation and further demonstrates the importance of the cognitive construct to driving.

3.3 Discussion and conclusions

In the literature review, it was contended that the concept of SA has not been well defined in the context of driving and that there is an increasing need to understand the implications of in-vehicle automation and devices on driver SA. Based on existing SA theory, this pilot study developed an operational definition of SA in the driving domain and applied it to a medium fidelity simulation to provide further insight into the importance of interaction with in-vehicle systems to human perception, comprehension and projection of states of the driving environment. Specifically, the effects of ACC and cell phone use on driver SA, workload and driving task performance were assessed, and the study described the extent to which secondary tasks compete for driver mental recourses.

In general, the results of the pilot study provide support for the application of in-vehicle automation, like ACC, under normal driving conditions for facilitating driver SA.

It appears that ACC control relieves drivers of vehicle monitoring and motor control workload, and they may pay more attention to the driving environment (as a primary task). Consequently, drivers may develop more complete and accurate knowledge of driving states (SA). It is possible that this benefit of automation to driving SA (under normal driving conditions) may lead to observed improvements in overall performance. That is, as all levels of SA improve in driving, there appears to be links to improvements in operational driving behaviors, as suggested in the transactional model of SA, in this context. These inferences differ from those of Ward (2000), who observed reduced driver SA through performance data under hazardous driving circumstances (e.g., responses to unexpected pedestrian crossings). However, the positive performance implications of ACC, which we observed, are in agreement with all prior work, including Parker et al. (2003), particularly improvements in variation in headway distance and following speed control.

This study also provided further support for the hypothesis that (hand-held) cell phone usage can be detrimental to driver SA (Gugerty et al., 2003). The experiment provided evidence that cell phone conversations (as a secondary task during driving) compete for limited mental resources of drivers. Consequently, drivers may not pay enough attention to the driving environment (as a primary task) and they may not develop complete and accurate knowledge of driving vehicle states (SA). This decrease in SA may lead to decrements in driving performance. Cell phone use, like in-vehicle automation use, appears to be an underlying factor in the linkages of the levels of SA in driving (save perception) to operational behaviors under normal conditions. These inferences are in agreement with the findings of Gugerty et al. (2003) on driving SA effects of cell phone use. Although Chen and Lin (2003) and Hancock et al. (1999, 2002) demonstrated significant driving performance decrements due to cell phone use (e.g., missed braking responses), we did not observe similar effects with our freeway simulation of a following task under normal conditions. We did expect performance decrements due to the cell phone use, but the short period of the cell phone conversations during our experiment may not have been sufficient to cause problems. Cell phone conversations may result in significant deleterious effects on driving performance with longer, continuous conversations.

In this study, only high-level driver SA appeared to be sensitive to the interaction effect of in-vehicle automation and device use. Similar to Rudin-Brown et al. (2003) results, we found that the benefits of ACC, in terms of workload reduction, were offset by workload increases due to cell phone use, driver distraction from the primary task, and associated degradations in SA. It is possible that the negative impact of the interaction of these technologies may be more pervasive across the levels of SA (perception, comprehension and projection) under more complex, interactive driving conditions posing higher mental workload. However, the current study only supported sensitivity of Level 3 SA to the ACC and cell phone interaction condition and, therefore, a linkage of driving environment state prediction to operational task performance in the transactional model of SA in driving.

Figure 11 presents an update of the transactional model of SA based on the findings of the pilot study. The labels "ACC" and "Cell" on the lines in the graph identify the mediating effects of the ACC and/or cell phone on SA for specific types of driving behavior. As mentioned before, the solid lines represent a critical link, and the dashed lines represent a potential or weak link, between SA and driving task types in the graph.

On the basis of this pilot study, directions of future research include developing broader operational definitions of SA in driving, which apply to more than the freeway following tasks examined here, as well as additional empirical work to identify other in-vehicle system factors that may be influential in driver SA. More specifically, there is a need to study the interaction effect of in-vehicle automation and device use, like navigation aids and cell phones, on driver SA in complex navigation, hazard negotiation or emergency driving conditions using direct, objective measures of the construct.

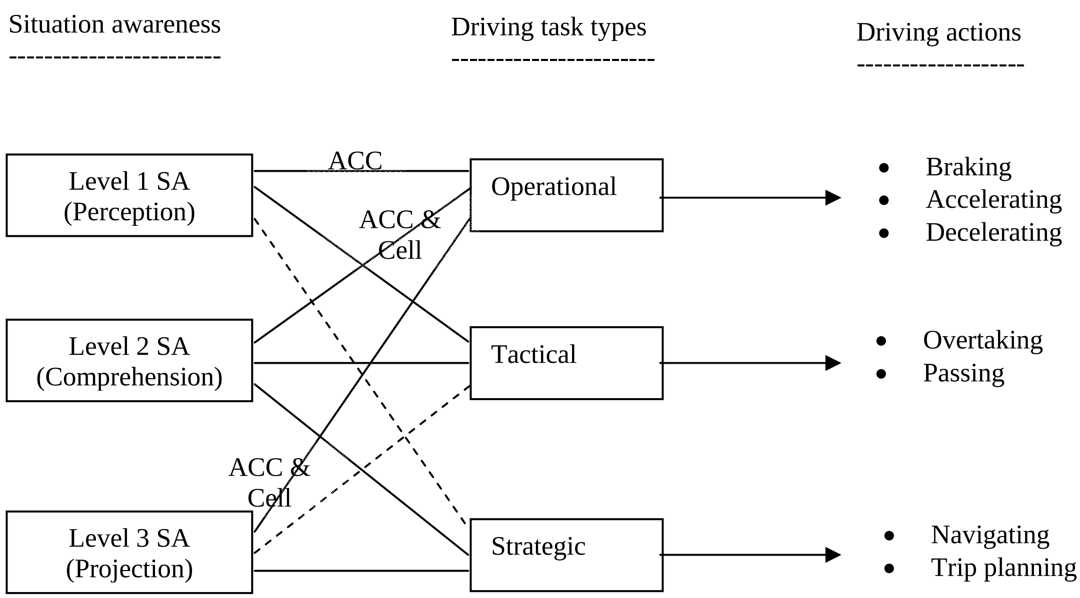


Figure 11: Updated transactional model of situation awareness and its potential influences on task types and driver actions

Another direction of future work, closely related to the present study, would be to investigate the impact of advanced automation LKS to determine if the effects on driver SA, workload and performance are comparable to those of the ACC system. Furthermore, it would be worthwhile to examine the compound effect of using multiple forms of in-vehicle automation on driver SA when confronted with secondary tasks, like cell phone use. These directions of work are beyond the scope of the current research but represent important studies towards advancing understanding of how SA functions in a broad range of driving contexts.

The mental resources of drivers will continue to be stretched in the future by the advent of new-sophisticated in-vehicle automation and more elaborate portable, personal communication and data assistance devices used while driving. There is a need to continue to investigate how drivers will achieve and maintain SA in the presence of this technology in support of safe driving performance. Resulting knowledge should be applied to the development of future technologies, or the redesign of existing devices, and reflected in any state and/or federal regulations on in-vehicle device use. In general, future research efforts should be focused on increasing driver SA under normal driving circumstances to better negotiate highway systems and to be prepared for emergency events.

4 Experimental Methodology

4.1 Objectives

The study presented in this section extended the pilot study by examining a more complex driving navigation task, the use of in-vehicle navigation assistance systems, as well as the source and reliability of assistance, as potential underlying factors in the linkages of SA to operational and strategic driving task performance. The study also sought to identify any effects of the in-vehicle automation and driving task factors on driver trust and workload responses. The overarching goal of the study was to further support and refine the transactional model of levels of SA in driving task types and specific driver behavior presented in Figures 11.

4.2 Task

The driving simulation used in the pilot study was enhanced for this study. Participants drove a virtual car and performed a freeway driving task, and a navigation task. The freeway driving simulation was similar to the pilot study simulation. The driving navigation simulation presented a model of a suburban area, including five street blocks adjacent to the existing freeway loop.

The viewpoint of the simulation provided to drivers was also from inside a virtual car. All roadways were marked with conventional lines. There were also different types of traffic signs along the suburb area streets, including: “pedestrian crossing”, “slow”, “deer crossing”, “railroad”, “speed limit”, and “stop” signs. There were street name signs in the suburban area. Beyond signs, the new suburban environment extension included office buildings, grass, rivers and streetlights.

The complete simulation was naturally divided into the two sub-tasks (freeway driving and suburb navigation). In the first task, participants were required to drive on the freeway, using ACC or manual control mode, of which they were informed in advance. Participants were instructed to keep their vehicle in one of the right-hand freeway lanes and maintain a speed of 60 mph. They were also asked to observe all road signs. Near the end of this task, participants received navigation information from one of two sources: a human aid or in-vehicle automation aid at 9 minutes into the trial. Participants interacted with the human aid by using a cell-phone with a head set. They alternatively interacted with the automation aid by viewing a laptop display screen, separate from the driving simulation screen. Participants were requested to make a response when the navigation aid was activated. The human or automation aids instructed the driver on the time required to reach the freeway exit and begin the navigation task. The freeway driving task was terminated just before participants reached the freeway exit.

Participants were subsequently required to drive through the suburban area obeying all traffic signs and following the navigation aiding information from the human or automation aid without deviating (see script and display examples for the human and automation aid in Appendix C) from the directions in their driving. The sports truck was manually controlled in the suburb navigation task. The navigation task required drivers to follow the guidance of aids on turning at certain streets and speed limits in order to safely turning at certain streets reach a specific destination (a “red” building) in the suburban area. Figure 12 presents the suburb navigation area with street names and the destination, which was marked by a red circle. The destination was located four (suburban) blocks away from the freeway exit and participants were required to make five or more turns to reach the building. In all navigation trials, participants had a paper map at their disposal, which showed the suburb area. Participants could see the destination in the upper right corner of the map and they could use the map to assess the efficiency and accuracy of route information provided by the navigation aids. The polar direction (north) was also shown on the map.

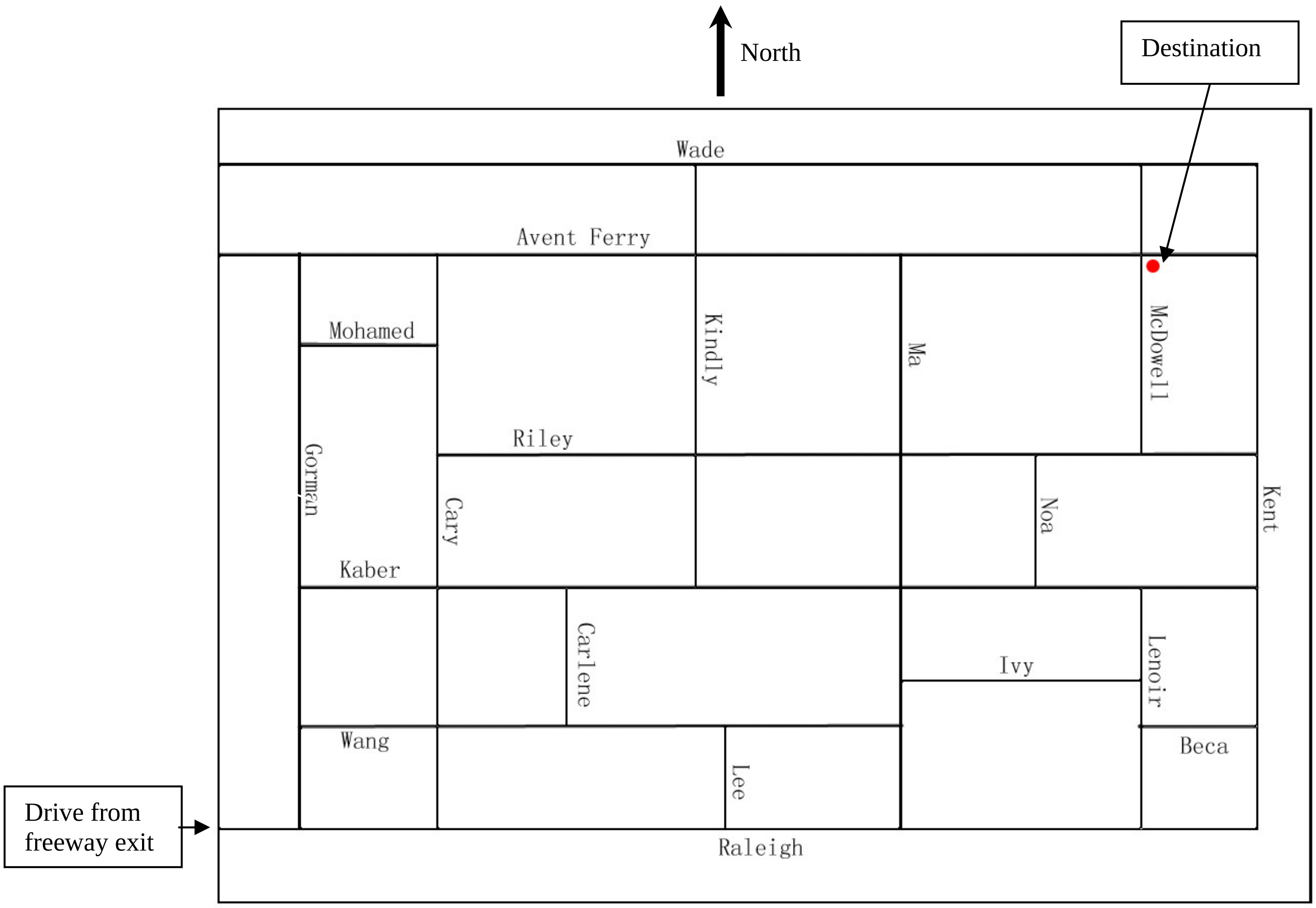


Figure 12: Paper map of the suburb area with street names and destination

4.3 Variables

4.3.1 Independent

The independent variables in the study included: (1) the use of in-vehicle automation or cruise control mode (i.e., ACC or manual control during the first freeway driving task - Experiment A); (2) the navigation information aid type (i.e., human-aiding via cell phone or automation-aiding via laptop display during the freeway driving and navigation tasks – Experiments A and B); and (3) the level of navigation aid reliability (during Experiment B), including 100%, 80% and 60%, as well as a control condition involving a telemarketing survey delivered through the navigation aid (see an example of the telemarketing script in Appendix D). The navigation aiding provided drivers with turning information (street names) and speed limit information. For example, the human aid would say (or the automaton aid would display a message), “Now, turn LEFT onto Cary Rd., and your driving speed should be increased to 45 mph”. All of the turning and speed limit information were correct according to the street signage, but the overall efficiency of the route information varied across reliability conditions in terms of the number of turns and the total elapsed driving time. The perfect, or 100% reliable, navigation condition required drivers to make 5 turns, and they spent an average of 11.7 minutes in reaching the destination from the freeway exit. The 80% reliable navigation condition required drivers to make 6 turns and an average navigation time of 14.2 minutes. The 60% reliable navigation condition required 7 turns and an average driving time of 15.1 minutes. The telemarketing survey was communicated by the human through the cell phone or by the automation aid through the laptop display. Participants were required to answer the survey verbally while an experimenter recorded their responses in writing. This served as a secondary distracter task during the driving navigation of the suburban area with the map and served to demonstrate any benefit of the navigation aids providing driving relevant information.

4.3.2 Dependent

Several dependent variables were observed during each experiment (or task).

(1) In Experiment A, driving task performance was measured in terms of lane maintenance deviations and consistency of speed control on the freeway (relative to the speed limit of 60 mph). Driver response time and accuracy in detecting and responding to information from the navigation aids was measured. The navigation information was activated at 9 minutes into each trial. Participates were required to click a button on the simulator steering wheel once they detected the information, and the computer recorded the time. The time elapsed was used as the driver response time. This served to quantify the impact of ACC use of performance with the navigation aid.

(2) During Experiment B, navigation task performance was measured in terms of driver adherence to the advice of the navigation aid (i.e., whether drivers followed the navigation aid in making correct turns), lane maintenance deviations, and consistency of speed control in the suburb area (relative to the posted speed limits). At the outset of each trial, drivers were provided with the map of the suburban area. As previously mentioned, the driving destination was revealed on the map. During the navigation task, drivers were instructed to follow the advice of the navigation aid regardless of whether they knew the route information was “incorrect”, or not the most efficient. Errors in following the

navigation aid were recorded along with time-to-task completion (which was dictated by the reliability of the aid information and participant motor control behavior).

(3) Driver SA was assessed at the end of each experiment (or task) using an adaptation of the SAGAT methodology (Endsley, 1995b). The SA questionnaire was similar to those employed in the pilot study; however, the pool of queries was expanded to address driver goals and decisions as part of the navigation task. For the freeway-driving task, participants were required to recall the locations of cars they passed, traffic sign information, and instructions from the navigation aid. They were required to identify any necessary driving behaviors (acceleration, braking and turning) in following the instructions of the aids in exiting the freeway (when, where, how fast) (see the questionnaires in Appendix E). In the navigation task, participants were asked to recall the drive time between turns, and when they passed certain signs or streets. They were also asked to recall the drive time to the destination and possible optimal solutions to reach the destination. For example, one SA query asked, "What route would have generated the shortest drive time to reach the destination when you passed the intersection of Kaber St. and Ma St.?" There was no time limit on participant responding to SA queries. After participants completed the SA questionnaire (freeze) at the close of the freeway-driving task, they returned to the driving simulation and continued performance of the navigation task. Example of the SA questionnaire for the present experiment is shown in Appendix E. As in the pilot study, all SA queries were presented on paper and participants responded in writing with pencil.

(4) Driver workload was also collected at the close of each experiment. The NASA- Task Load Index (TLX) (Hart & Staveland, 1988) was used to capture participant

perceptions of physical, mental, temporal, effort, frustration and performance demands at the end of the freeway driving task (Experiment A) and at the end of the navigation task (Experiment B). Participants were provided with a sheet of demand factor descriptions. They then completed a subjective demand ranking form requiring pairwise comparisons of all demand factors, and they identified the demands they believed would more greatly affect performance in the experimental tasks. (These rankings were made for both freeway driving and suburb navigation, following participant training in the simulation.) Following each task, participants completed a subjective rating of the workload demand factors. They were required to draw vertical lines on linear scales for each of the demand factors (there were six linear scales on a rating form) at the position they felt best represented the demands for a specific test trial. (See Appendix F for demand ranking and rating form.) In order to obtain a composite workload score for each participant, the ratings were multiplied by weighting factors calculated based on the demand component rankings. A composite workload score (from 0 to 100) was obtained for each trial in each experiment.

(5) Driver trust in the navigation aids was measured using a subjective survey of initial participant trust expectations as well as a subjective rating at the close of each trial as part of Experiment B. Participants completed a trust evaluation of the automation aid or human aid at the beginning of the study on the basis of prior personal experiences and expectations. They rated how well they thought the aid would perform and how many errors they thought it would make. Additional trust ratings were collected at the end of each subsequent test trial involving human or automation navigation aiding. An adaptation of the survey form used by Dzindolet et al. (2002) was employed for the initial

expectations assessment (see Appendices G and H) along with a trust rating form for the multiple successive navigation trials (see Appendices I and J).

4.4 Experimental designs

Experiment A was a two-factor experiment, including two levels of cruise control (ACC and manual/no-ACC) and two levels of navigation aid type (human and automation aid). Both variables were manipulated between-subjects. An equal number of participants (10 people) were randomly assigned to the two levels of the cruise control. Half of each of these groups was randomly assigned to human or automation aiding, as shown in Table 1. There was only one trial for each participant during the experiment (participants were considered repeated measures on the conditions) and the human and automation aids provided 100% reliable driving directions during that trial (i.e., freeway exit information).

Table 1: Data collection table based on design of Experiment A

	Cruise control mode			
	Manual / No-ACC		ACC	
	Navigation source type		Navigation source type	
	Human aid	Automation aid	Human aid	Automation aid
Participant Number	1,2,3,4,5	6,7,8,9,10	11,12,13,14,15	16,17,18,19,20

A mixed 2 x 4 design was used for Experiment B, including two levels of navigation source type (human aid and automation aid) and four levels of navigation aid reliability. The navigation source type was manipulated as a between-subjects variable. The navigation aid reliability was controlled as a within-subjects variable. Half of the participants who completed Experiment A under each of the cruise control conditions (manual/No-ACC or ACC) were assigned to the human-aid condition in Experiment B. The remaining participants from Experiment A were assigned to the automation-aid condition. The data collection table for Experiment B is presented in Table 2. (The participant numbers can be matched across Tables 1 and 2.)

Table 2: Data collection table based on design of Experiment B

<table>
<tr><td colspan="2" rowspan="3">Within-Subject Variable</td><td colspan="2">Between-Subject Variable</td></tr>
<tr><td colspan="2">Navigation source type</td></tr>
<tr><td>Human aid</td><td>Automation aid</td></tr>
<tr><td rowspan="4">Navigation reliability</td><td>100% reliable</td><td rowspan="4">Participant Number:

1,2,3,4,5,

11,12,13,14,15</td><td rowspan="4">Participant Number:

6,7,8,9,10,

16,17,18,19,20</td></tr>
<tr><td>80% reliable</td></tr>
<tr><td>60% reliable</td></tr>
<tr><td>Task-irrelevant info and map</td></tr>
</table>

Directly following the completion of Experiment A (the freeway drive), each participant completed three successive navigation task trials involving the various navigation aid reliability conditions (100%, 80% and 60% reliable) and an additional trial involving the telemarketing survey and map use. The reliability conditions were presented to participants in decreasing order. The participants assigned to human or automation aiding followed the same trial orders beginning with the 100% reliable condition and finishing with the control condition. That is, the experimental condition setup was the same for the two levels of the navigation source type. This approach to delivery of the experiment was necessary to examine the change in driver trust in the two types of aids over time, as the reliability of each aid degraded.

4.5 Apparatus

Figure 13 presents the experiment equipment setup. The same equipment that was used in the pilot study was used in this experiment. The stereo display emitter (a StereoGraphics CE2) and Goggles were used for viewing the 3-D display. Drivers input control actions through the realistic steering wheel, and gas and brake pedals. A Motorola T720 cell phone was used by the drivers to receive navigation information from the human aid. An IBM ThinkPad R31 laptop computer was used to present navigation information from the automation aid. A copy of the suburb map was taped to the desk, adjacent to the VR driving simulation display. Drivers could refer to a map of the suburb area for the driving navigation task at any time.

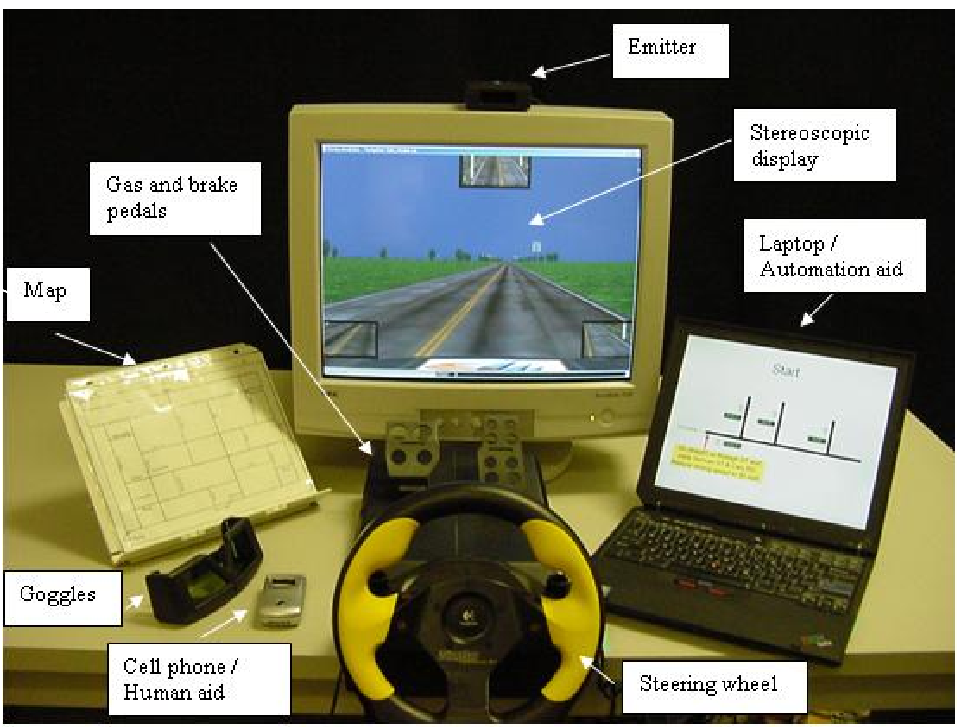


Figure 13: Experiment setup

4.6 Participants

Twenty participants from the NCSU undergraduate and graduate student populations were used in this study. Recruits were required to have 20/20 or corrected to normal vision, and to have at least three years of driving experience in order to participate. These recruitment criteria were intended to ensure adequate visual performance in the task and to prevent learning affects on simulation performance because of a potential lack of participant driving experience.

4.7 Procedure

Table 3 presents a list of all the steps in the experimental procedure along with average times for each. The total experiment time for each participant was approximately 3 hours and 30 minutes.

Table 3: Overview of experimental procedure and approximate time estimates

Step in procedure	Time (minutes)
1. An Introduction to the study, including informed consent (see Appendix K).	20
2. Collection of anthropometric survey data (see Appendix L).	5
3. Training in the simulation driving tasks (freeway driving and suburb navigation) under a manual control mode and without aids.	10
4. Familiarization with the SAGAT method and questionnaire, as well as the NASA-TLX workload rating scale to be administrated at the close of trials as part of Experiments A and B (see Appendices E and F).	15
5. Familiarization with the initial trust expectations survey and ratings forms (see Appendices G to H) and the post-trial trust survey and ratings forms (see Appendices I to J).	5
6. One complete training trial (freeway driving and suburb navigation), including the SA questionnaires and workload ratings (manual control and no aiding).	25

7. One test trial as part of Experiment A to assess the ACC effect on navigation aid use, including a SA questionnaire and workload rating at the close.	25
8. Four consecutive navigation task trials, including SA questionnaires and workload ratings with intervening 5-min breaks between trials.	100

Before participants began the test trials, they were provided with a dedicated training session. They were provided with training on how to control the virtual car in the driving simulator and how to maintain the vehicle on the virtual roadway using the physical steering wheel, gas and brake pedals. Participants experienced both freeway driving and suburb navigation under manual control of the car and without any navigation aiding (as shown in Step 3 of the procedure). There were no traffic signs or street names presented in the simulation environment for this training session. Participants were also provided with a second training session (as shown in Step 6 of the procedure) in which the simulation remained the same. Participants were requested to answer example SA questions at the end of the freeway driving and suburb navigation. Drivers always had access to the paper navigation map, and they were informed of the location of the driving destination during the training session. This second training experience was intended to account for potential learning effects across the reliability conditions (the within-subjects variable settings) in Experiment B.

During Experiment A, the navigation aid was turned on 9 min into the trial. Participants reached the freeway exit at approximately 11 min into the trial and the simulation was terminated. Just prior to this, the simulation was temporarily frozen, and the SA questionnaire on the freeway driving and environment was administrated. Subsequently the TLX workload rating was administered (see Figure 14 for a schematic of all trial events). The duration of the cell phone call from the human aid is also presented in Figure 14.

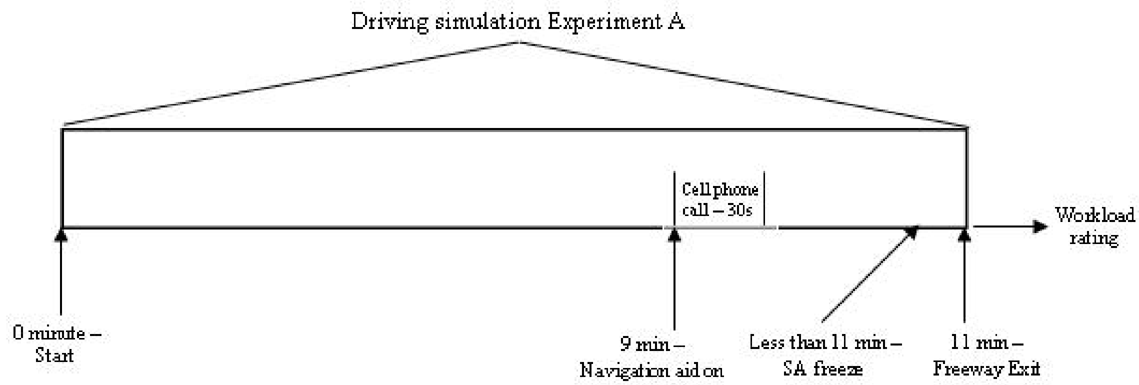


Figure 14: Schedule of events in Experiment A

During Experiment B, participants initially completed the trust expectation survey on the automation or human aid. The ratings as part of this survey served as baseline measures of human trust in the aids. Additional trust ratings were collected at the end of each subsequent test trial involving human or automation aiding. Participants also completed a SA questionnaire and NASA-TLX workload rating at the end of each trial (see Figure 15 for a schedule of experiment trial/events). During the navigation task, the aids presented information to drivers as they approached each street intersection or decision point. A “wizard-of-oz” technique was used in which an experimenter, observing test participants driving performance on a separate remote monitor, called the participant on the cell phone, or the experimenter controlled the automated navigation aid computer to deliver specific driving direction displays to participants at the “right” time. Drivers always had access to the paper map, and an optimal route for navigation was only marked on the map under the control (telemarketing survey) condition. The duration of the cell phone calls and breaks between each call from the human aid are also presented in Figure 15. There were total of 6 to 8 cell phone calls and 5 to 7 breaks (depending upon the navigation aid reliability) during the course of the suburb navigation driving. Scripts of the verbiage spoken to participants at each intersection were written in advance of Experiment B for each human aid reliability condition. Similarly, static navigation aid display screens were developed for delivery of different driving instructions to participants for each automation aid reliability condition (see Appendix C for examples).

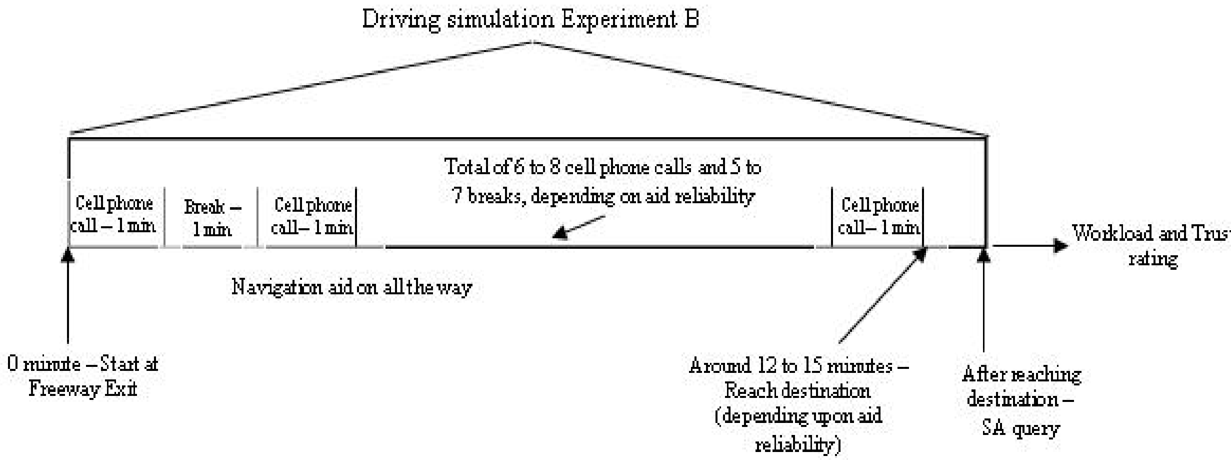


Figure 15: Schedule for navigation aiding and SA queries during Experiment B

5 Hypotheses

5.1 Task performance

Based on the results of the pilot study, in Experiment A the ACC system was expected to facilitate better speed control than the manual condition because of the potential for driver vigilance decrements over extended periods of manual control. It was also expected that driver detection (speed and accuracy) of information from the navigation aid during the freeway-driving task would be better when the ACC was active, as a result of workload relief provided by the automation. It was expected there would be comparable effects of the automation aid and the human aid on driving performance. The cell phone conversation with the human aid was expected to distract driver attention and the additional automated aid visual display was expected to be equally distracting to freeway driving.

Based on the results of the pilot study, in Experiment B, it was also expected that presentation of task-irrelevant information via the cell phone or the automation aid display during the navigation task would significantly degrade driver performance (navigation errors, time-to-task completion), as compared to driving with the human or automation aid providing task-relevant information. Higher reliability navigation aiding was expected to facilitate better driving performance in Experiment B.

5.2 Workload

Based on the pilot study, in Experiment A it was generally expected that the use of the ACC system would reduce driver workload, as compared to the manual condition, which required participants to monitor for, and implement, speed changes in freeway driving. This study used the NASA-TLX as a measure of driving workload. The measure was expected to reveal many aspects of workload in driving, including physical and mental demands.

For both Experiments A and B, since an additional visual attention/load was to occur during navigation performance with the automation aid, and driving is primarily a visual-motor task, according to multiple resource theory (Wickens, 1984), it was expected that there would be a higher perceived workload for the trials with automation aiding than trials involving human aiding. This was because the driving task and automation navigation aiding were expected to pose similar perceptual demands on drivers.

Concerning the different levels of aiding reliability in Experiment B, human operators were expected to pay more attention to the lower reliability aid conditions as a result of inaccurate or "puzzling" driving guidance being presented; consequently, it was expected that there would be higher perceived workload than when using higher reliability navigation aiding. The task irrelevant information presentation condition was expected to distract driver attention from the primary driving task and significantly increase the perception of workload, beyond all other conditions involving the various forms of navigation aiding.

5.3 Situation awareness

Based on the results of the pilot study, in Experiment A it was expected that use of the ACC system would facilitate improvements in driver SA by reducing task load in terms of the need to monitor for, and implement, vehicle speed changes. Attentional resources would be freed-up for perceiving, for example, roadway signage and relating this information to driving goals.

During the navigation task (Experiment B), since the cell phone conversation with the human navigation aid was expected to distract driver attention from the driving environment, this condition was also expected to reduce SA, when navigation information was unreliable. Similarly, since driving is basically a visual and motor control process, the visual search demands associated with retrieving information from the automation-aid display were expected to be substantial and to reduce SA on the driving environment, when navigation information was unreliable. In the perfect reliability condition, it was possible that the information provided by the aid would benefit driver SA and out-weight any decrements in driving environment perception due to in-vehicle device distractions. It was expected that presentation of task-irrelevant information via the cell phone or automated aid would significantly degrade driver SA during navigation performance, as compared to driving with a navigation aid providing task relevant information.

The various levels of human and automation aid reliability were also expected to have a main effect on driver SA. People were expected to perceive the need to pay more attention to the navigation information being presented under low reliability settings and make comparison with their own judgments on the driving situation, based on references to the suburban area map. In general, it was expected that there would be higher SA scores with 100% reliable guidance from the human advisor or automation aid than with 80%, and 60% reliable navigation aiding.

5.4 Trust

It was hypothesized that participants would initially say that they trusted the automation aid more than the human advisor, since people have expectation (based on daily experiences) that automation is generally reliable and may relieve them of some mental and physical workload (Dzindolet et al., 2002). However, based on Wiegmann et al., (2001) and Dzindolet et al. (2001) work, when participants experienced automation failures/errors during Experiment B, it was expected that their trust would decline more sharply than trust in the human advisor. That is, there may be an interaction effect of information aid sources and in-vehicle automation reliability on trust, since people may trust in the human aid and automation aid in different ways (Dzindolet et al., 2001).

6 Data Analyses

All statistical analyses were performed using SAS (Statistical Analysis Software). They included multivariate analyses of variance (MANOVAs) and multi-way analyses of variance (ANOVAs) applied to the different dependent variables. MANOVAs were conducted on the response measures observed during Experiments A and B for which inter-correlations were expected, including the collections of performance measures and SA measures. ANOVAs were then conducted on all performance and SA variables for those significant main effects and interactions revealed through MANOVA results. ANOVAs were also conducted on workload and trust dependent variables, but these variables were not included in the MANOVAs as there was no *a priori* expectation of inter-correlations among workload and trust, etc. When a significant interaction effect was found based on ANOVA results, additional simple effects analyses were conducted to further establish/confirm the main effects of independent variables on the responses. The full ANOVA model for Experiment A can be written as follows:

$$Y_{ijr} = \mu + A_i + N_j + A*N_{ij} + \varepsilon_{ijr}$$

where,

Y_{ijr} = the response variable (e.g., speed deviation, workload, SA);

A: ACC control mode;

N: navigation aid type;

i = 1, 2;

j = 1, 2;

r = 1, 2, 3, 4, 5.

SAS PROC GLM was used for analyzing the statistical model of Experiment A. An alpha level of 0.05 was used to identify any significant main effects of A or N and the presence of any significant interaction effect. Further investigation of significant predictors was conducted using Duncan's Multiple Range tests with an alpha criterion of 0.05.

The full ANOVA model for Experiment B can be written as follows:

$$Y_{ijk} = \mu + R_i + N_j + Sub(N)_{k(j)} + R*N_{ij} + \varepsilon_{ijk}$$

where,

Y_{ijk} = the response variable (e.g., driving error, workload, SA and trust);

R: navigation reliability levels;

N: navigation aid type;

Sub (N): subject (nested in N);

i = 1, 2, 3, 4;

j = 1, 2;

k = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

SAS PROC MIXED was used to analyze the statistical model for Experiment B due to the nature of the experiment design. The design was mixed with between- and within-subjects variables in a split-plot layout. The statistical model included multiple error terms that are appropriately handled by PROC MIXED for F-tests and post-hoc procedures. An alpha level of 0.05 was used to identify any significant main effects of N or R and the presence of any significant interaction effect. Further investigation of significant predictors or an interaction was conducted using Tukey's tests with an alpha criterion of 0.05.

For both statistical models for Experiments A and B, residual analyses were conducted to ensure that the underlying assumptions of normality and constant variance of the ANOVA were upheld by the data sets. Residual plots, normal probability plots, and normality statistics (Shapiro-Wilks test) were used to verify these conditions. The outcomes of the participants initial trust expectation for Experiment B failed to meet the normality assumptions of the ANOVA. (The test results were significant, but the data set did not conform to the assumptions of the analysis.) In light of this, and the discrete nature of the trust rating data, the initial trust expectation observations were subjected to nonparametric analyses, based on ranks. The Kruskal-Wallis test was used to determine if there was a significant main effect of the navigation aid type.

Correlation analyses were conducted to identify any potential relationships among the various response measures recorded during the experiments, including: (1) task performance measures and the SA measures; (2) SA measures and the subjective measure of workload; (3) SA measures and the subjective measure of trust; (4) subjective workload measures and the subjective measure of trust; and (5) task performance and the subjective measure of trust. Pearson Product-Moment coefficients were calculated to establish the strength of any positive or negative linear associations of the responses. The SAS PROC CORR procedure was used to establish the statistical significance of the correlations of interest to this study.

7 Results

7.1 Participant characteristics

The average age of the participants was 28.1 years with a standard deviation of 5.1 years. All persons had 20/20, or corrected to normal, vision. As part of the anthropometric data survey, participants were asked to rate their prior experience with VR applications, in playing video games, simply using a PC, driving, using control while driving, using a cell phone while driving, using navigation assistance while driving, and using a map while driving. They were also asked how many years they have had been driving. With respect to VR experience, the average response (on a scale from 1 = "none" to 5 = "frequent") was low (1.9 with a standard deviation of 0.9). With respect to playing video games, on average participants indicated moderate experience (2.9 with a standard deviation of 1.1). With respect to PC experience, the average participant rating indicated a high frequency of use (4.8 with a standard deviation of 0.7). With respect to driving frequency, the average response was high (4.9 with a standard deviation of 0.3). With respect to experience in cruise control use, on average participants indicated moderate experience (2.6 with a standard deviation of 1.4). With respect to cell phone use while driving, on average participants indicated moderate experience (2.6 with a standard deviation of 1.1). With respect to navigation assistance experience while driving, the average response was low (1.2 with a standard deviation of 0.5). With respect to using a map while driving, the average participant rating indicated a moderate frequency (2.4 with a standard deviation of 0.9). Finally, in regard to years of driving experience, the average for all participants was 8.5 years with a standard deviation of 4.5.

7.2 Experiment A

7.2.1 Task performance

MANOVA results revealed significant effects of the ACC control mode ($F(2, 15)=79.41$, $p<0.0001$) and navigation aid type ($F(2,15)=4.73$, $p=0.0255$) on the collection of performance measures, including speed deviations and response time to navigation information. There was no significant interaction effect of the ACC control mode and navigation aid type.

7.2.1.1 Speed control

ANOVA results revealed a significant effect of the ACC control mode on variations in driver speed control ($F(1,19)=147.71$, $p<0.0001$). There proved to be no main effect of the navigation aid type for this response measure. Figure 16 presents the RMSE of speed control for the ACC control mode (mean deviation of participant vehicle speed from posted limits). The ACC control mode and navigation aid type condition means and standard deviations for speed control deviations are included in Table 4. There were significantly greater deviations in speed with manual control.

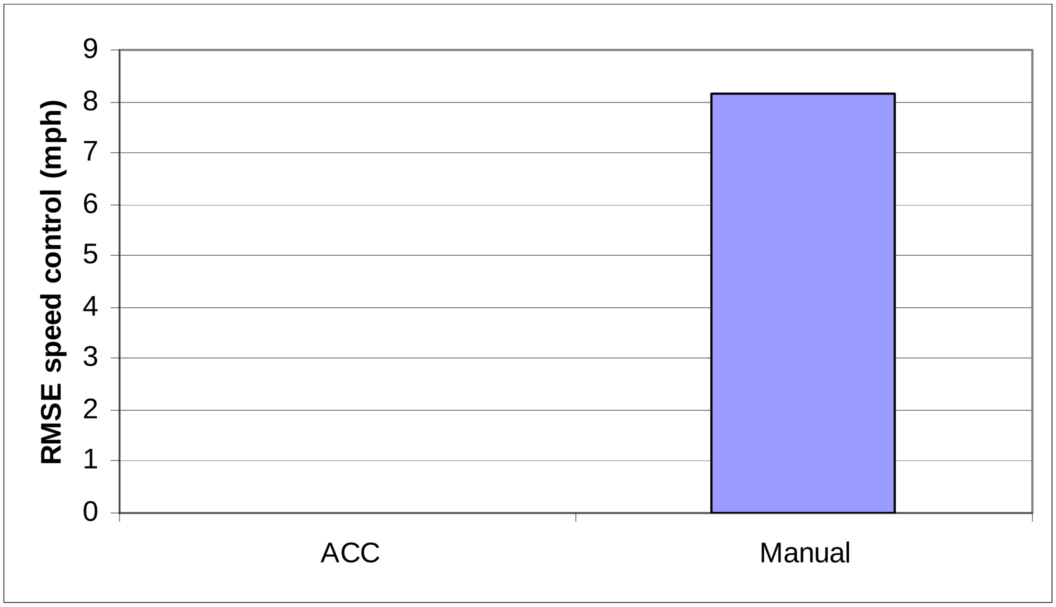


Figure 16: RMSE of speed control for ACC control mode

7.2.1.2 Response time (to navigation aiding)

ANOVA results revealed significant effects of the ACC control mode ($F(1,19)=4.57$, $p=0.0484$) and navigation aid type ($F(1,19)=8.49$, $p=0.0101$) on response time to detect the navigation information presented by either the human or automated aid. Figure 17 presents the average response time for ACC control mode and navigation aid type conditions. The ACC control mode and navigation aid type condition means and standard deviations for driver response time are also included in Table 4. The response time was significantly longer when the ACC control was inactive. There was also a significantly longer driver response time to the presentation of navigation information when the automation aid was being used. The relevant MANOVA and ANOVA results on driving performance measures for Experiment A are summarized in Table 5.

Table 4: Means and standard deviations (in parentheses) for driving performance, SA and workload measures for all ACC control mode and navigation aid type settings for Experiment A

Independent variables		Dependent variables									
		Driving performance		Situation awareness				Workload			
		Speed deviations (mph)	Response time to navigation info (s)	Overall SA	Level 1 SA	Level 2 SA	Level 3 SA	Overall TLX	Physical	Frustration	Effort
ACC control mode	ACC	0 (0)	2.1 (1.5)	0.77 (0.14)	0.73 (0.31)	0.67 (0.22)	0.90 (0.16)	0.44 (0.17)	0.22 (0.26)	0.21 (0.19)	0.31 (0.23)
	Manual	8.14 (2.05)	4.2 (3.4)	0.52 (0.11)	0.43 (0.22)	0.43 (0.16)	0.70 (0.25)	0.64 (0.17)	0.53 (0.31)	0.48 (0.28)	0.68 (0.23)
Navigation aid type	Automation aid	4.30 (4.79)	4.6 (3.3)	0.62 (0.17)	0.57 (0.27)	0.57 (0.22)	0.73 (0.21)	0.53 (0.20)	0.42 (0.35)	0.34 (0.26)	0.47 (0.29)
	Human aid	3.86 (4.23)	1.7 (0.8)	0.67 (0.19)	0.60 (0.34)	0.53 (0.17)	0.87 (0.23)	0.55 (0.20)	0.34 (0.31)	0.36 (0.30)	0.52 (0.32)

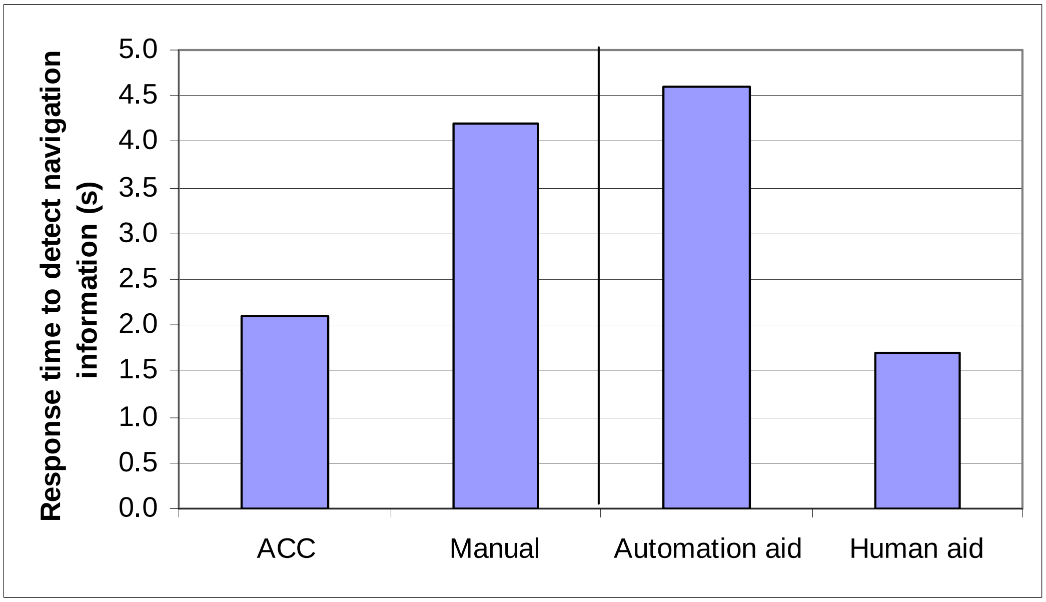


Figure 17: Response time for ACC control mode and navigation aid type

Table 5: Relevant MANOVA and ANOVA results on driving performance measures for Experiment A

Independent variables	MANOVA (Wilks' lambda)	ANOVA results: Dependent variables – Speed deviations	ANOVA results: Dependent variables – Response time to navigation information
ACC control mode	F(2,15)=79.41, p<0.0001	F(1,19)=147.71, p<0.0001	F(1,19)=4.57, p=0.0484
Navigation aid type	F(2,15)=4.73, p=0.0255	F(1,19)=0.43, p=0.5192	F(1,19)=8.49, p=0.0101
ACC control mode *Navigation aid type	F(2,15)=1.03, p=0.3826	N/A	N/A

7.2.2 Driver SA

MANOVA results revealed a significant effect of the ACC control mode (F(4,13)=5.65, p=0.0074) on SA measures. There was no significant main effect of the navigation aid type and no interaction effect of the ACC control mode and navigation aid type on the SA measures.

Figure 18 presents the mean Level 1, Level 2, Level 3 and overall SA scores for ACC control mode conditions. The ACC control mode and navigation aid type condition means and standard deviations for the SA scores are also included in Table 4. ANOVA results on driver SA indicated that the ACC control mode was influential in the percentage of correct responses to SA queries on perception of the driving environment, comprehension and projection. There was a significant effect of ACC control mode on Level 1 SA ($F(1,19)=6.97$, $p=0.0178$), Level 2 SA ($F(1,19)=7.00$, $p=0.0176$), Level 3 SA ($F(1,19)=4.79$, $p=0.0438$), and overall SA ($F(1,19)=18.45$, $p=0.0006$). In general, Figure 18 reveals that the mean accuracy for SA was higher when the ACC system was active. The relevant MANOVA and ANOVA results on driver SA measures for Experiment A are presented in Table 6.

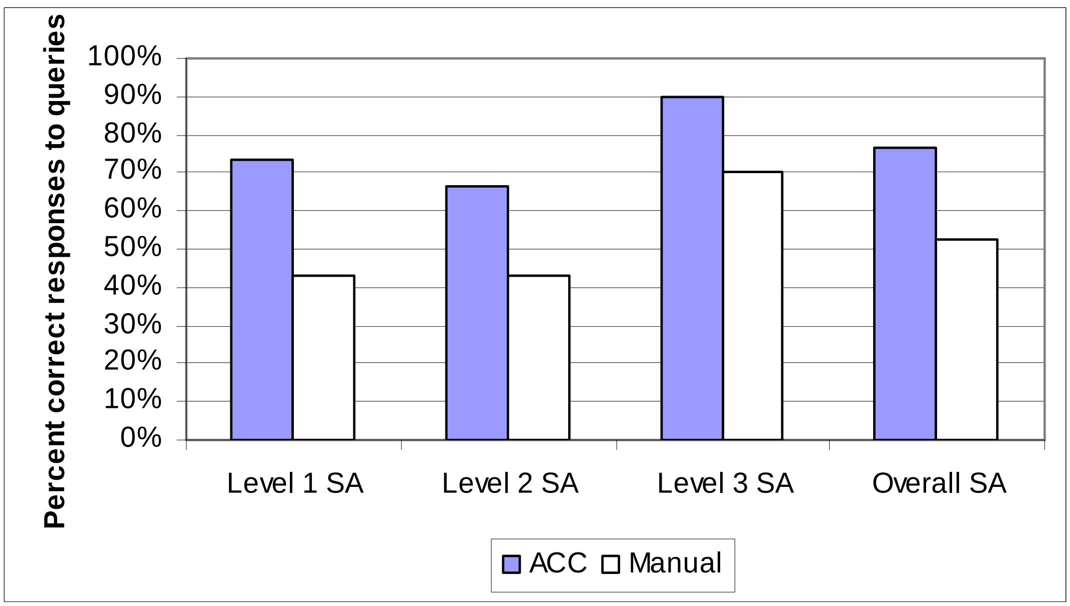


Figure 18: Mean percent correct responses to SA queries for ACC control mode

Table 6: Relevant MANOVA and ANOVA results on driver SA for Experiment A

Independent variables	MANOVA (Wilks' lambda)	ANOVA results			
		Dependent variables			
		Overall SA	Level 1 SA	Level 2 SA	Level 3 SA
ACC control mode	F(4,13)=5.65, p=0.0074	F(1,19)=18.45, p=0.0006	F(1,19)=6.97, p=0.0178	F(1,19)=7.00, p=0.0176	F(1,19)=4.79, p=0.0438
Navigation aid type	F(4,13)=2.14, p=0.1330	N/A	N/A	N/A	N/A
ACC control mode *Navigation aid type	F(4,13)=1.99, p=0.1556	N/A	N/A	N/A	N/A

7.2.3 Driving workload

Figure 19 presents the mean NASA-TLX scores (overall and individual demand factors) for the ACC control mode. The ACC control mode and navigation aid type condition means and standard deviations for the workload ratings (including the overall TLX score and any demand component for which there was an observed significant effect) are also included in Table 4. ANOVA results revealed significant effects of the ACC system on the overall TLX score (F(1,19)=6.80, p=0.0191), the physical demand rating (F(1,19)=5.45, p=0.0329), frustration ratings (F(1,19)=5.67, p=0.0300), and effort ratings (F(1,19)=12.05, p=0.0031). There was no main effect of the navigation aid type and no interaction effect of the ACC control mode and navigation aid type on perceived workload. The ANOVA results on driver workload, including overall TLX scores and all of the demand components for Experiment A are summarized in Table 7.

In agreement with the hypothesis on perceived workload, Figure 19 reveals significantly higher ($p<0.05$) workload scores when the ACC system was inactive. It is important to note that the scale for performance ratings, as part of the NASA-TLX, is reversed in comparison to the scales for all other demand factors. Consequently, the plot in Figure 19 reveals that, on average, participants thought they performed better when using the ACC.

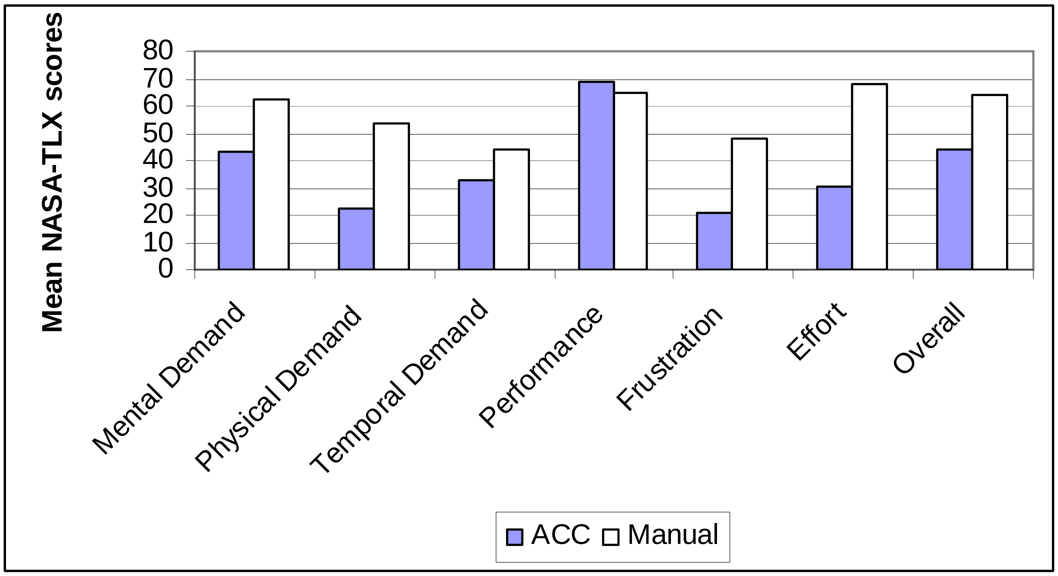


Figure 19: Mean NASA-TLX scores for ACC control mode

Table 7: F-test results on driver workload for Experiment A

Independent variables	Dependent variables						
	Overall TLX	Mental	Physical	Temporal	Performance	Frustration	Effort
ACC control mode	F(1,19)=6.80, p=0.0191**	F(1,19)=2.40, p=0.1407	F(1,19)=5.45, p=0.0329**	F(1,19)=0.69, p=0.4185	F(1,19)=0.33, p=0.5748	F(1,19)=5.67, p=0.03**	F(1,19)=12.05, p=0.0031**
Navigation aid type	F(1,19)=0.12, p=0.7358	F(1,19)=0.72, p=0.4096	F(1,19)=0.32, p=0.5798	F(1,19)=0.07, p=0.7891	F(1,19)=0.43, p=0.5221	F(1,19)=0.03, p=0.8622	F(1,19)=0.24, p=0.6303
ACC control mode *Navigation aid type	F(1,19)=0.10, p=0.7546	F(1,19)=0.03, p=0.8751	F(1,19)=0.01, p=0.93	F(1,19)=0.12, p=0.7343	F(1,19)=0.96, p=0.3409	F(1,19)=0.02, p=0.8896	F(1,19)=0.13, p=0.7228

** -- significant at p<0.05 level.

7.2.4 Correlation analyses

Simple correlation analyses were conducted in order to identify any significant relationships among driving performance, SA and workload for the experiment. A

Pearson correlation coefficient revealed significant negative linear associations between overall SA and variations in speed control (r = -0.6628, p=0.0014). There was also a significant negative linear association between Level 1 and Level 2 SA and variations in speed control (r = -0.5012, p=0.0244 and r = -0.6010, p=0.0051, respectively). All these findings indicate positive associations of the constructs of SA and driving performance; that is, as SA increased, the RMSE in speed control decreased.

Pearson correlation coefficients also revealed significant negative linear associations between the overall TLX rating and the overall SA score (r = -0.6431, p=0.0022), Level 1 SA (r = -0.6476, p=0.0020), and Level 2 SA (r = -0.5634, p=0.0097). That is, as perceived workload increased, driver ability to perceive and comprehend the driving environment decreased. Beyond this, there were significant negative linear associations between the physical demand factor of the TLX and the overall SA score (r = -0.5335, p=0.0154), the Level 1 SA score (r = -0.4469, p=0.0482), and the Level 2 SA score (r = -0.6324, p=0.0028). The frustration factor was negatively correlated with the overall SA score (r = -0.6888, p=0.0008), the Level 1 SA score (r = -0.5340, p=0.0153), and the Level 2 SA score (r = -0.6266, p=0.0031). Similarly, the effort factor was negatively correlated with the overall SA score (r = -0.7141, p=0.0004), the Level 1 SA score (r = -0.6338, p=0.0027), and the Level 2 SA score (r = -0.4627, p=0.04). All of these correlations indicate that driver ability to develop SA on the driving environment decreased when the specific factors of the TLX (such as physical demand, frustration and effort) increased. There was no significant correlation of perceived workload with Level 3 SA. This may be attributable to the nature of the freeway driving, which imposed a lower cognitive load, in general, as there were few operational behavior requirements (i.e., speed changes, turns, etc.). This is consistent with the comparatively weak effect of ACC control on Level 3 SA score (i.e., p=0.0438), as revealed by the ANOVA.

Finally, Pearson correlation coefficients revealed significant linear associations between the variation in speed control and the overall TLX rating (r = 0.5857, p=0.0067), the physical demand factor of the TLX (r = 0.5911, p=0.0061), the frustration factor (r = 0.5008, p=0.0245), and the effort factor (r = 0.6441, p=0.0022). That is, as perceived workload increased, driver performance decreased (with higher speed control variation).

7.3 Experiment B

7.3.1 Task performance

MANOVA results revealed a significant effect of the navigation aid reliability (F(6,106)=3.99, p=0.0012) on the collection of performance measures, including speed deviations and driving errors. There was no significant main effect of the navigation aid type and no interaction effect of the navigation aid type and aid reliability on the collection of performance measures.

7.3.1.1 Speed control

ANOVA results revealed a significant effect of navigation aid reliability on variations in driver speed control (F(3,79)=3.49, p=0.0217). Figure 20 presents the RMSE of speed control for the navigation reliability conditions. The navigation aid type and aid reliability condition means and standard deviations for speed control deviations are included in Table 8. Tukey’s test revealed significantly greater deviations in speed (p<0.05) when participants performed under the control condition (i.e., telemarketing survey) as compared to the 80% and 60% reliable aiding conditions.

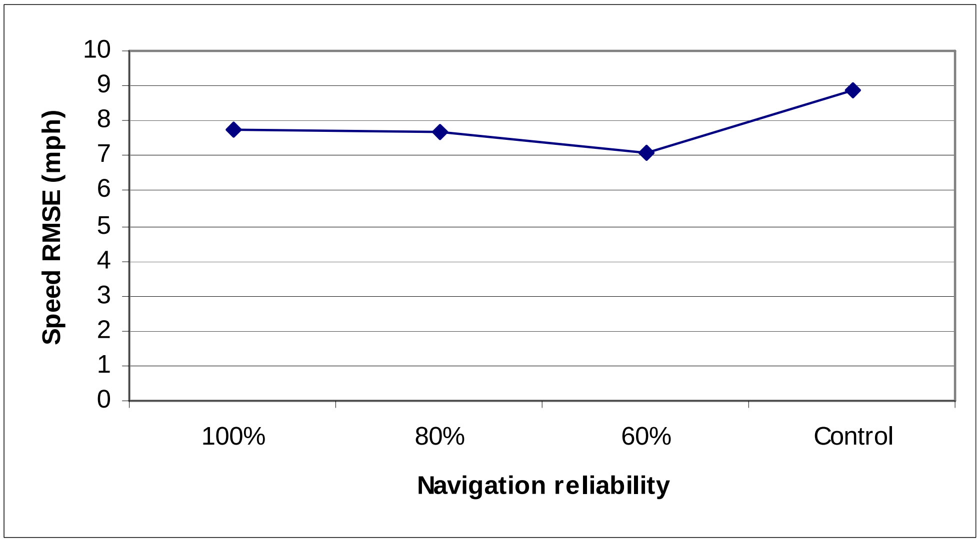


Figure 20: RMSE of speed control for navigation aid reliability conditions

7.3.1.2 Driving errors

ANOVA results revealed a significant effect of navigation aid reliability on driving errors (deviations from directions) ($F(3,79)=5.00$, $p=0.0039$). Figure 21 presents the driving errors across the various navigation aid reliability conditions. The navigation aid type and aid reliability condition means and standard deviations for the driving errors are also included in Table 8. Tukey's test revealed significantly lower driving errors ($p<0.05$) when participants received 100% reliable aiding as compared to all other conditions. In general, participants made more driving errors when the aiding was more inefficient and in the control condition (telemarketing survey). The relevant MANOVA and ANOVA results on driving performance for Experiment B are summarized in Table 9.

Table 8: Means and standard deviations (in parentheses) for driving performance, SA and workload measures for all navigation aid type and aid reliability settings for Experiment B

Independent variables		Dependent variables									
		Driving performance		Situation awareness				Workload			
		Speed deviations (mph)	Driving errors	Overall SA	Level 1 SA	Level 2 SA	Level 3 SA	Overall TLX	Mental	Temporal	Frustration
Navigation aid type	Automation aid	8.54 (3.38)	0.43 (0.59)	0.53 (0.14)	0.57 (0.28)	0.52 (0.23)	0.51 (0.25)	0.62 (0.18)	0.63 (0.25)	0.59 (0.22)	0.60 (0.21)
	Human aid	7.17 (2.14)	0.58 (0.64)	0.49 (0.14)	0.48 (0.31)	0.52 (0.28)	0.48 (0.23)	0.61 (0.19)	0.65 (0.26)	0.45 (0.26)	0.41 (0.23)
Navigation aid reliability	100%	7.77 (2.41)	0.15 (0.37)	0.55 (0.13)	0.32 (0.30)	0.75 (0.18)	0.57 (0.25)	0.62 (0.19)	0.62 (0.26)	0.50 (0.24)	0.50 (0.24)
	80%	7.66 (2.44)	0.50 (0.61)	0.53 (0.12)	0.55 (0.25)	0.53 (0.20)	0.52 (0.23)	0.61 (0.18)	0.64 (0.23)	0.47 (0.25)	0.46 (0.24)
	60%	7.11 (2.04)	0.60 (0.68)	0.54 (0.13)	0.70 (0.26)	0.38 (0.25)	0.52 (0.17)	0.58 (0.18)	0.58 (0.27)	0.51 (0.25)	0.56 (0.22)
	Control	8.88 (4.13)	0.75 (0.64)	0.43 (0.15)	0.52 (0.26)	0.40 (0.21)	0.38 (0.27)	0.64 (0.19)	0.71 (0.25)	0.60 (0.27)	0.49 (0.26)

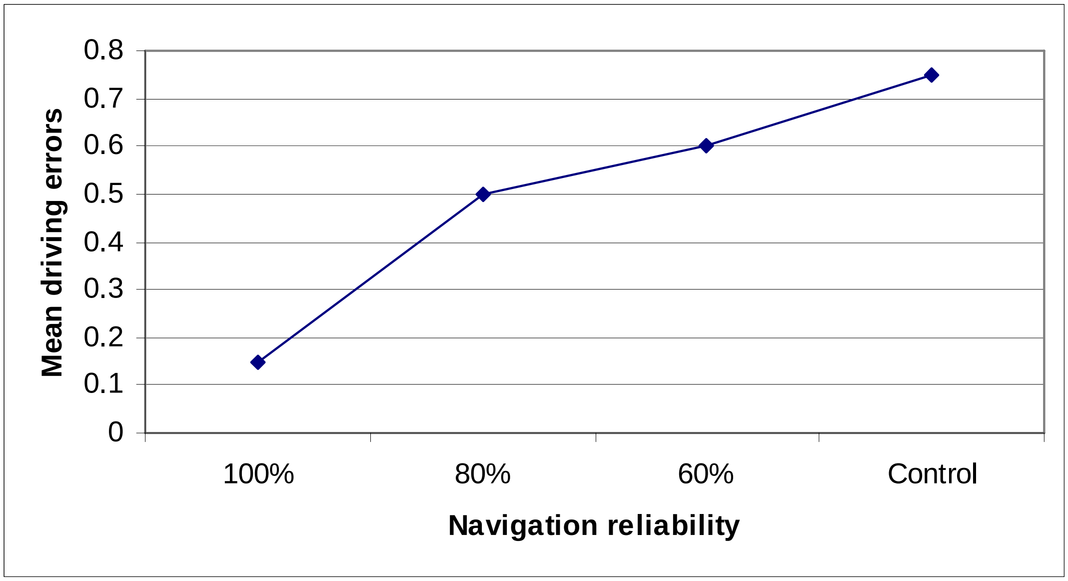


Figure 21: Driving errors across various navigation aid reliability conditions

Table 9: Relevant MANOVA and ANOVA results on driving performance measures for Experiment B

Independent variables	MANOVA (Wilks' lambda)	ANOVA results Dependent variables Speed deviations	ANOVA results Dependent variables Driving errors
Navigation aid reliability	F(6,106)=3.99, p=0.0012	F(3,79)=1.77, p=0.0217	F(3,79)=5.00, p=0.0039
Navigation aid type	F(2,17)=1.62, p=0.2275	N/A	N/A
Navigation aid type *Aid reliability	F(6,106)=1.85, p=0.0955	N/A	N/A

7.3.2 Driver SA

MANOVA results revealed a significant effect (F(12,135)=5.01, p<0.0001) of the navigation aid reliability on SA measures. There was no significant main effect of the navigation aid type and no interaction effect of the navigation aid type and aid reliability on the SA measures.

Figure 22 presents the mean Level 1, Level 2, Level 3 and overall SA scores for the various navigation aid reliability conditions. The navigation aid type and aid reliability condition means and standard deviations for the SA scores are also included in Table 8. The plot reveals that, on average, drivers exhibited better overall SA when aid information was 100% reliable and the worst SA when there was no navigation aiding provided and participants were required to address the telemarketing survey. Exceptions included Level 1 SA for the 100% reliable condition and Level 2 SA for the 60% reliable condition. In general, the reliability factor significantly affected SA at all levels. Each response measure is addressed below.

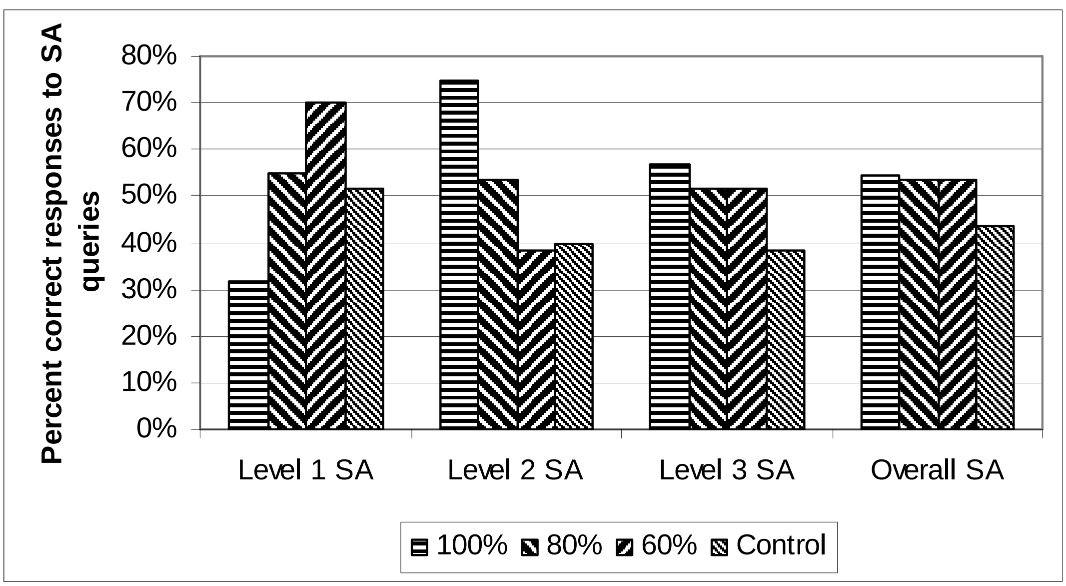


Figure 22: Mean percent correct responses to SA queries for navigation aid reliability conditions

ANOVA results revealed a significant effect of navigation aid reliability ($F(3,79)=7.19$, $p=0.0004$) on the percent correct responses to Level 1 SA queries. Tukey's tests were conducted to further investigate this main effect. In agreement with hypothesis, the post-hoc procedure revealed significantly higher ($p<0.05$) perceptual knowledge of the driving environment with the 60% and 80% reliable aiding as compared to the control condition; however, the worst perceptual knowledge occurred with the 100% reliable aiding condition. It is possible that since drivers realized the human advisor or automated aid was highly accurate in direction under the latter condition, they simply did not pay as much attention to observing aspects of the driving environment.

ANOVA results revealed a significant effect of navigation aid reliability ($F(3,79)=16.78$, $p<0.0001$) on Level 2 SA. According to Tukey's tests, the 100% reliable navigation aiding produced significantly greater ($p<0.05$) percent correct responses to Level 2 SA queries in comparison to all other conditions. The 80% navigation aid reliability also produced significantly greater ($p<0.05$) percent correct responses to Level 2 SA queries, as compared to the control condition and the 60% navigation aid reliability condition.

ANOVA results revealed a significant effect of navigation aid reliability ($F(3,79)=3.17$, $p=0.0314$) on Level 3 SA (projection of states of the driving environment). Tukey's test revealed significantly higher ($p<0.05$) Level 3 SA scores for all navigation aid reliability levels, as compared to the control condition (i.e., any navigation aiding was better than none in terms of Level 3 SA).

ANOVA results also revealed a significant effect of navigation aid reliability (F(3,79)=4.05, p=0.0114) on overall SA. According to Tukey's tests, the navigation aiding trials with 100%, 80% and 60% reliable task information produced significantly greater (p<0.05) percent correct responses to overall SA queries in comparison to the control condition. The relevant MANOVA and ANOVA results on driver SA measures for Experiment B are summarized in Table 10.

Table 10: Relevant MANOVA and ANOVA results on driver SA measures for Experiment B

Independent variables	MANOVA (Wilks' lambda)	ANOVA results			
		Dependent variables			
		Overall SA	Level 1 SA	Level 2 SA	Level 3 SA
Navigation aid reliability	F(12,135)=5.01, p<0.0001	F(3,79)=4.05, p=0.0114	F(3,79)=7.19, p=0.0004	F(3,79)=16.78, P<0.0001	F(3,79)=3.17, p=0.0314
Navigation aid type	F(4,15)=0.59, p=0.6724	N/A	N/A	N/A	N/A
Navigation aid type *Aid reliability	F(12,135)=0.92, p=0.5338	N/A	N/A	N/A	N/A

7.3.3 Workload

Figure 23 presents the mean NASA-TLX scores (overall and individual demand factors) for the navigation aid reliability condition. The navigation aid type and aid reliability condition means and standard deviations for the workload ratings (including the overall TLX score and any demand components for which there was an observed significant effect) are also included in Table 8. ANOVA results revealed significant effects of navigation aid reliability on driver perceived mental demand (F(3,79)=4.63, p=0.0059) and temporal demand (F(3,79)=7.11, p=0.0004). ANOVA results also revealed a significant effect of the navigation aid type on the frustration demand factor (F(1,79)=6.13, p=0.0235). There was no significant navigation aid type by aid reliability interaction effect. The ANOVA results on driver workload, including the overall TLX score and all of the demand components for Experiment B are summarized in Table 11.

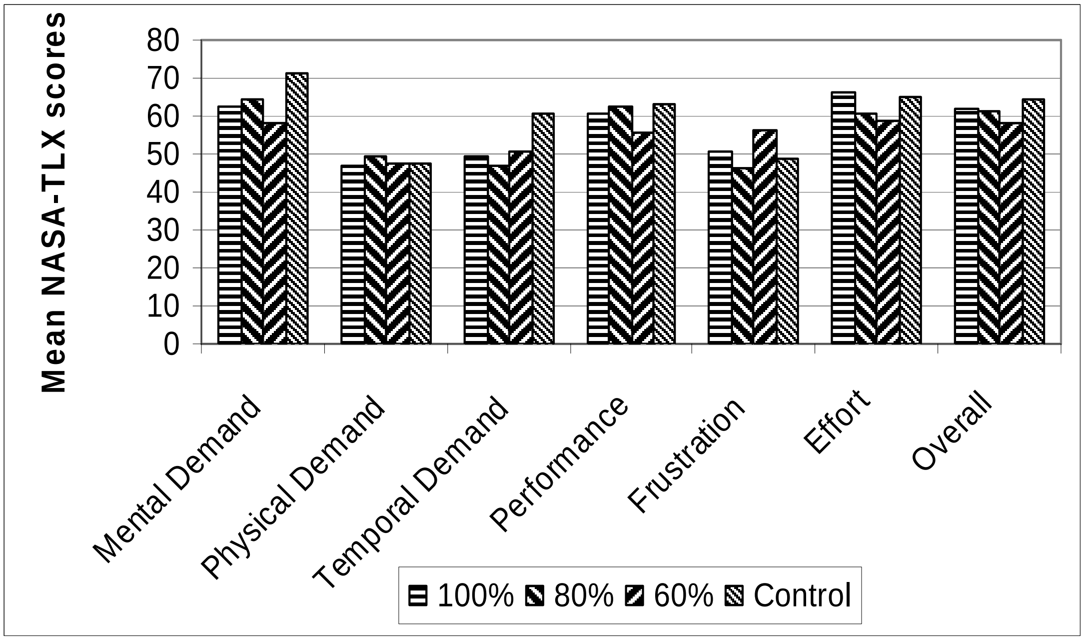


Figure 23: Mean NASA-TLX scores for navigation aid reliability

Table 11: F-test results on driver workload for Experiment B

Independent variables	Dependent variables						
	Overall workload	Mental	Physical	Temporal	Performance	Frustration	Effort
Navigation aid reliability	F(3,79)=2.51, p=0.0681	F(3,79)=4.63, p=0.0059**	F(3,79)=0.53, p=0.6605	F(3,79)=7.11, p=0.0004**	F(3,79)=1.24, p=0.3049	F(3,79)=1.18, p=0.3264	F(3,79)=2.14, p=0.1054
Navigation aid type	F(3,79)=0.04, p=0.8414	F(3,79)=0.03, p=0.8713	F(3,79)=0.44, p=0.5170	F(3,79)=1.93, p=0.1822	F(3,79)=1.62, p=0.2193	F(3,79)=6.13, p=0.0235**	F(3,79)=0.18, p=0.6764
Navigation aid type *Aid reliability	F(3,79)=0.65, p=0.5880	F(3,79)=1.74, p=0.1701	F(3,79)=0.28, p=0.8389	F(3,79)=1.50, p=0.2240	F(3,79)=0.17, p=0.9187	F(3,79)=0.36, p=0.7791	F(3,79)=1.24, p=0.3025

** -- significant at p<0.05 level.

Tukey's tests on the NASA-TLX scores were conducted to further investigate the significant navigation aid type and reliability main effects. In partial agreement with the hypothesis on workload, the post-hoc test revealed significantly higher ($p<0.05$) frustration scores when the automation aid was used. This may be attributable to the additional visual attention required by the aid, distracting from the primary driving task (a visual-motor task). Tukey's test also revealed significantly higher ($p<0.05$) mental and temporal demand scores for the telemarketing survey, as compared to any navigation aiding whatsoever (60, 80 or 100%). Task-relevant information was likely easier for participants to process while driving than the task-irrelevant information.

7.3.4 Trust

Figure 24 presents the initial expected mean scores (initial trust expectation and initial expected errors) for the navigation aid types. The navigation aid type and aid reliability condition means and standard deviations for the initial trust ratings are included in Table 12. Since the initial trust expectation and expected error data sets violated the ANOVA normality assumption, nonparametric analyses of the data were conducted to investigate whether there was a significant effect of the navigation aid type. Kruskal-Wallis tests revealed significantly higher initial trust expectations ($T(1)=8.38$, $p=0.0038$), and fewer expected errors ($T(1)=13.07$, $p=0.0003$) for the automation aid, as compared to the human aid. These results were in-line with expectations.

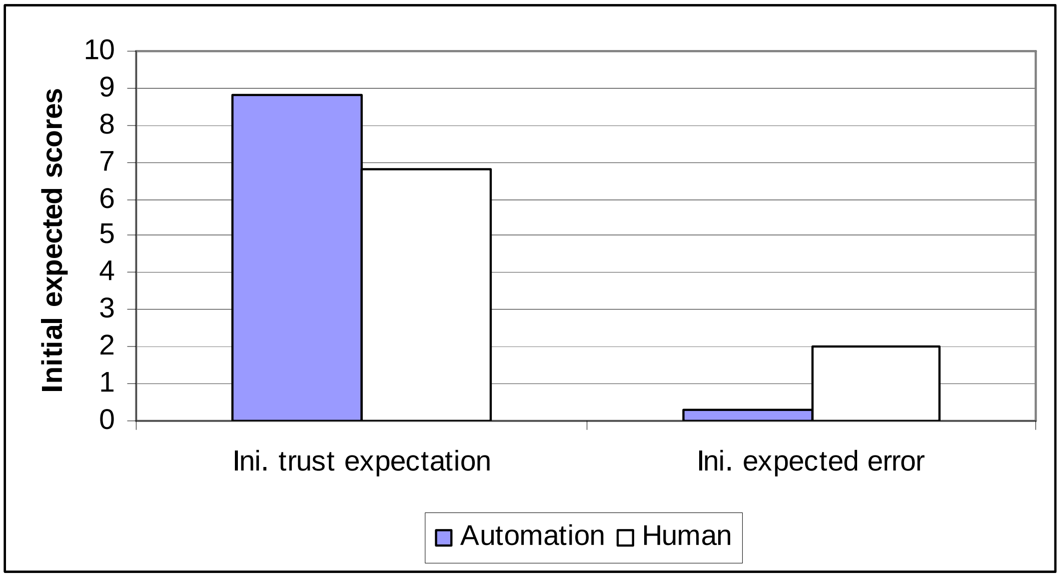


Figure 24: Initial expected mean scores for navigation aid type

Table 12: Means and standard deviations (in parentheses) for trust ratings for all navigation aid type and aid reliability settings

Independent variables		Dependent variables		
		Initial trust expectation	Initial expected errors	Post-trial trust ratings
Navigation aid type	Automation aid	8.8 (1.0)	0.3 (0.5)	6.83 (1.98)
	Human aid	6.8 (1.5)	2.0 (1.2)	6.91 (1.96)
Navigation aid reliability	100%	N/A		8.25 (1.20)
	80%			6.80 (1.57)
	60%			5.58 (2.05)
	Control			N/A

Figure 25 presents the mean trust expectation scores for each aid type and the interaction effect of the aid type and reliability on all of the trust ratings. The navigation aid type and aid reliability condition means and standard deviations for the post-trial trust ratings are also included in Table 12. The composite of the trust rating data did not violate the normality assumption of the ANOVA and parametric F-test. ANOVA results revealed a significant effect of navigation aid reliability ($F(3,79)=19.80$, $p<0.0001$) on trust rating scores. According to Tukey's tests, the 100% reliable condition and initial trust expectation produced significantly higher ($p<0.05$) ratings than the 80% and 60% reliable conditions. Tukey's test also revealed significantly higher ($p<0.05$) trust rating scores when participants received 80% reliable aiding compared to the 60% reliable condition.

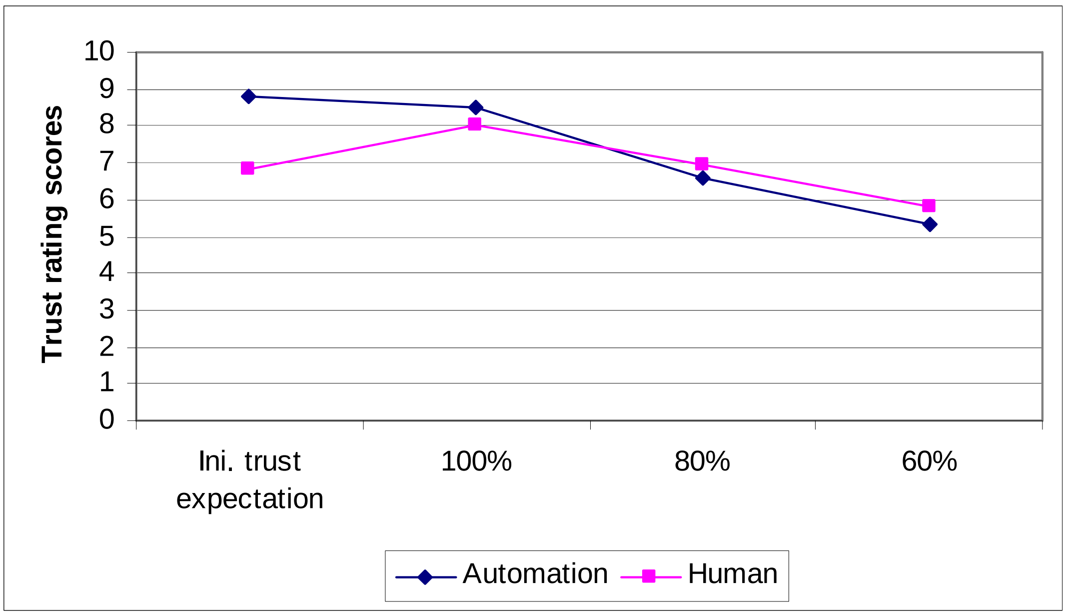


Figure 25: Navigation aid type and reliability interaction effect on trust

ANOVA results also revealed a significant interaction effect of the navigation aid type and reliability settings on all trust ratings ($F(3,79)=4.42$, $p=0.0075$). The results of Tukey's test ($p<0.05$) on the interaction effect are shown in Table 13. The conditions are sorted in the table from highest to lowest trust rating. The results of this analysis indicated that participant initial trust expectations for automation aid were significantly different from all trust ratings for imperfect navigation aiding excluding the 80% reliable human aid. For imperfect navigation aiding conditions (80% and 60% reliable), on average, participants assigned higher trust ratings to the human than the automation aid. The 60% reliable automation aid led to the worst trust ratings and it was significantly worse than perfect human and automation aiding. These findings are in agreement with the hypotheses on the trends of trust for the various aiding conditions and reliabilities.

Simple effects analyses were also conducted to explore the interaction effect in more detail. Results supported the main effects of navigation aid reliability ($p<0.05$) and aid type ($p<0.05$) on the trust rating scores for the complete trust rating data set.

Table 13: Results of Tukey's test on the navigation aid type and reliability interaction for trust ratings

Tukey-Kramer Grouping			Mean Trust Ratings	Navigation Reliability	Navigation Aid Type
A			8.8	Initial rating	Automation
A			8.5	100%	Automation
A	B		8.0	100%	Human
A	B	C	6.95	80%	Human
A	B	C	6.8	Initial rating	Human
	B	C	6.65	80%	Automation
		C	5.8	60%	Human
		C	5.35	60%	Automation

An additional ANOVA was conducted on only the post-trial trust ratings, which did not reveal a significant interaction effect of the navigation aid type and aid reliability settings. There was a significant effect of the navigation aid reliability ($F(2,59)=26.35$, $p<0.0001$) on the post-trial ratings. According to Tukey's test, there were significant differences ($p<0.05$) in trust ratings among all three levels of aid reliability. There was no significant effect of the navigation aid type on the post-trial trust ratings. The ANOVA and Kruskal-Wallis test results on driver trust for Experiment B are summarized in Table 14.

Table 14: F-test and Kruskal-Wallis test results on driver trust for Experiment B

Independent variables	Dependent variables			
	Initial trust expectation	Initial expected errors	Post-trial trust (including initial ratings)	Post-trial trust (excluding initial ratings)
Navigation aid reliability	N/A	N/A	F(3,79)=19.80, p<0.0001**	F(2,59)=26.35, p<0.0001**
Navigation aid type	T(1)=8.38, p=0.0038**	T(1)=13.07, p=0.0003**	F(3,79)=0.67, p=0.4246	F(2,59)=0.02, p=0.8936
Navigation aid type * Aid reliability	N/A	N/A	F(3,79)=4.42, p=0.0075**	F(2,59)=0.96, p=0.3932

** -- significant at p<0.05 level.

7.3.5 Correlation analyses

Simple correlation analyses were conducted in order to identify any significant relationships among driving performance, SA, workload and trust for Experiment B. A Pearson correlation coefficient revealed a significant negative linear association between Level 3 SA scores and driving navigation errors ($r = -0.2464$, $p=0.0276$). This result indicates a positive association of the Level 3 SA construct with driving performance in terms of navigation; that is, as Level 3 SA increased, navigation errors decreased.

Finally, a Pearson correlation coefficient also revealed a significant linear association between the post-trust ratings and Level 2 SA ($r = 0.3485$, $p=0.0064$). That is, as drivers achieved higher comprehension of states of the driving environment, relative to task goals, they assigned higher ratings of trust to the navigation aids, in general.

8 Discussion

8.1 Driving performance

In the freeway driving task (Experiment A), the ACC system reduced the deviation in speed control, consistent with the results of the pilot study (Ma and Kaber, 2005). As expected, the ACC system also produced shorter response times for driver detection of navigation information from the human or automated aid. These findings suggest driver vigilance decrements may occur over extended periods of manual control. They also indicate automation may provide relief of driving workload allow for quicker responses to navigation information from aids. These observations are also supported by the higher TLX workload score for the manual driving condition, and the positive correlation between the variation in speed control and the overall TLX rating. That is, as driver workload increased speed control variation also increased.

In the suburb navigation driving (Experiment B), the higher reliability navigation aiding did not serve to reduce the deviation in speed control, which was hypothesized. This finding may be attributable to a trial order effect. The experiment presented the aiding reliability conditions in the order 100%, 80% and 60%, followed by the control condition (telemarketing survey). Drivers may have become more and more familiar with the suburb driving area (e.g., speed limit on the roads) across the reliability conditions and, therefore, reduced deviations in speed control occurred under lower reliability navigation aiding. However, the control condition (presenting task-irrelevant information) did produce the worst speed deviations. This finding demonstrates the benefit of navigation aiding providing driving relevant information, in general. It appeared that the automation aid produced worse speed control in the navigation driving task because of a conflict for visual attention resources with the primary driving task, particularly when the aid presented irrelevant (distracter) information under the control condition (i.e., telemarketing survey). However, there was no significant main effect of aid type on speed control.

In general, higher reliability navigation aiding facilitated better driving performance measured in terms of navigation errors. However, the automation aid and human aid appeared to cause comparable distraction effects on driver navigation performance. The automation aid demanded driver visual attention for perception of navigation information, but posed a lower memory requirement as the aid display was not transient like the human auditory message. This led to worse speed deviations but fewer driving errors. Although the human aiding did not demand the transfer of driver visual attention from the driving task, the directions from the human required more memory for storing verbal stimuli. This form of aiding supported speed control but not reductions in driving errors.

Although there was no difference among conditions in terms of deviations in lane maintenance, there was a trend for the ACC condition to produce fewer deviations in the freeway driving task. The control condition appeared to produce a greater mean deviation in the suburb navigation driving task than any of the navigation aiding conditions. However, there were no statistically significant differences among the various levels of the independent variables in terms of lane deviations across both experiments.

8.2 Driver workload

In the freeway driving task, the cruise control condition did not cause significant differences in perceived mental demands (TLX scores) for drivers. This is most likely because the freeway driving was basically a psychomotor activity that did not require significant cognitive resources. The following task investigated in the pilot study led to a significant difference in mental load (measured using a mental demand rating scale), which was attributable to the ACC system. This may suggest that car following requires more cognitive resources than only freeway driving in the absence of interactive traffic. This result is consistent with findings by Parker et al. (2003) and Rudin-Brown et al. (2003). The ACC systems they investigated also caused differences in various TLX demand components (physical demand, frustration and effort) when drivers were posed with more cognitively complex tasks than the lane and speed maintenance tasks as part of Experiment A. Taken together, these findings suggest that in-vehicle automation may relieve drivers of some workload in order to monitor, and implement, speed changes in more interactive freeway driving.

In the suburb navigation driving, there was no significant difference in overall workload (or TLX scores) attributable to the different navigation aid types. This was contrary to the hypothesis that the automation aid would lead to higher workload, based on multiple resource theory (Wickens, 1984). The automation and human aids appeared to produce comparable levels of workload in the driving task. However, there was a trend for the automation aid to produce higher overall mean workload ratings than the human aid for demand factors as part of the TLX, other than the mental demand. There was a trend for the human aid to pose greater mental demand than the automation aid. This may be attributable to the working memory requirement placed on drivers during the suburb navigation task to retain auditory driving directions. Although drivers were allowed to request repetition of the navigation information from the human aid, it appeared to be much easier for them to simply check the information display as part of the automation aid (but this was to the decrement of speed control).

The affects of the different levels of aiding reliability on perceived workload were mixed. There was no effect of navigation reliability on the overall TLX score. However, the control condition (telemarketing survey) produced significantly greater mental and temporal demands than the task-relevant information aiding conditions. This suggests the task-irrelevant information presentation condition distracted driver attention from the primary driving task and significantly increased the perception of workload. The different aid reliabilities, particularly the 80% and 60% conditions led to perception of higher cognitive workload. With respect to the other dimensions of the TLX, including overall effort, frustration or physical demands, there was no significant impact of in-vehicle automation reliability. The navigation (automation) aiding errors primarily affected the perception of cognitive aspects of workload.

8.3 SA in driving

The effects of ACC on SA in the freeway driving task (Experiment A) were consistent with the findings of the pilot study involving the lead car following task (Ma and Kaber, 2005). The ACC appeared to relieve drivers of some workload in terms of the need to monitor, and implement, vehicle speed changes. Drivers paid more attention to roadway signage and relating this information to driving goals. In a recent study by

Stanton and Young (2005), they found that ACC reduced driver SA, which was measured using the situation awareness rating technique (SART, by Taylor, Selcon and Swinden, 1995), in an overtaking and passing driving task. The task required participants to change/set the ACC mode. The authors argued that provision of a head-up display mirroring the ACC status from an instrument cluster display might reduce the SA decrements reported in their study. However, in the freeway driving and following tasks investigated in this research, the ACC was set at the beginning of experiment trials and drivers were not required to change the settings or monitor the status of the ACC. Of course, good driver SA or in-vehicle automation mode awareness may be more difficult to achieve if drivers are required to monitor and control additional in-vehicle status displays and, for example, to adjust low-speed and high-speed ranges for the ACC, according to various traffic and area conditions (Itoh, Inagaki, Shiraishi, Watanabe and Takae, 2005). How automation mode awareness in driving influences driver SA is still not entirely clear, and it likely depends on the nature of the driving task and how often drivers need to change/set system settings during the course of driving.

In the navigation driving task, the navigation aid type did not cause differences in driver perception, comprehension and projection of the driving environment. Driver distraction due to cell phone use (human aiding in the navigation) and visual attention allocation to the laptop visual display (automation aiding of the navigation) appeared to have comparable influences on driver's ability in terms of achieving SA. The pilot study and other research has provided evidence that cell phone use while driving degrades SA, particularly Level 2 and 3 SA (Ma and Kaber, 2005; Gugerty et al., 2003).

In the navigation driving task, the different levels of navigation aid reliability and the control condition (telemarketing survey) influenced driver ability in perception, comprehension and projection of the driving environment. As speculated in the Problem Statement, the varying reliability of in-vehicle navigation aiding appeared to attract driver attention away from the driving environment. Consequently, this had negative influence on driver SA, and possibly driving performance in terms of strategic driving behaviors. In general, the results of Experiment B revealed that higher navigation reliability produced higher driver SA (Level 2, 3 and overall SA). However, drivers achieved worst Level 1 SA in the navigation task under the highest aiding reliability condition. This finding may also attribute to a trial order effect. When drivers became more and more familiar with the suburb driving area, they were able to remember certain features, such as the road signs. Consequently, experience in the driving environment may have increased their accuracy of perception of the environment.

The manipulation of navigation aid reliability affected all levels (perception, comprehension and projection) of SA in the simulated navigation driving task. The navigation driving task was useful for demonstrating Level 1, Level 2, Level 3 and overall SA affects of the levels of navigation aid reliability and the control condition. Most importantly, the study confirmed linkages of strategic driving behavior to the three levels of SA. The transactional model of SA in driving, presented in Figure 11, summarized the relationships among the various levels of SA, the types of driving tasks/behaviors and relevant driving actions, based on the literature review (Matthews et al., 2001; Ward, 2000; and Endsley, 1995a) and empirical results of the pilot study (Ma and Kaber, 2005). Figure 26 presents an update of the transactional model of SA based on the findings of the navigation driving task (Experiment B). The solid and dashed linkages among the levels of SA and the tactical driving behaviors in the transactional model are based on other prior research. As mentioned before, the solid lines in the transactional model represent a critical link, and the dashed lines represent a potential or weak link between SA and driving task types in the diagram. According to Ward (2000), strategic driving behavior amounts to navigating and trip planning. In Experiment B, driver SA was significantly affected by navigation aid reliability and driver navigation performance was significantly correlated with changes in SA. The label “Aid Reliability” on the lines in Figure 26 identifies the mediating affect of the navigation aid reliability on SA for the specific type of driving behavior.

The current study, including driving navigation task performance, provided further insight into the importance of driver SA in strategic type driving tasks and associated actions. Specifically, varying the reliability of navigation aiding during trials caused differences in driver workload and SA and navigation knowledge. This, in turn, may have led to the corresponding changes in navigation performance. The updated transactional model of SA in driving provides more information on the nature and strength of association of specific aspects of SA with strategic driving behaviors. Contrary to expectation, the manipulation of the source of driver navigation information (aid type) did not produce variations in driver SA or serve as a further basis for describing linkages of levels of SA to strategic driving behaviors. The results of Experiment B demonstrated comparable influences of navigation aid type on driver ability in terms of achieving SA. Unfortunately, other investigations of the effects of in-vehicle automation (e.g., ACC) on driver SA, like the recent investigation by Stanton and

Young (2005), have not been able to confirm the potential linkages between levels of driver SA (perception, comprehension and projection) and tactical driver behavior/action like this study, because they have used indirect (performance-based) measures of SA or subjective rating techniques that do not assess each level of the theoretical construct, as the SAGAT measure does.

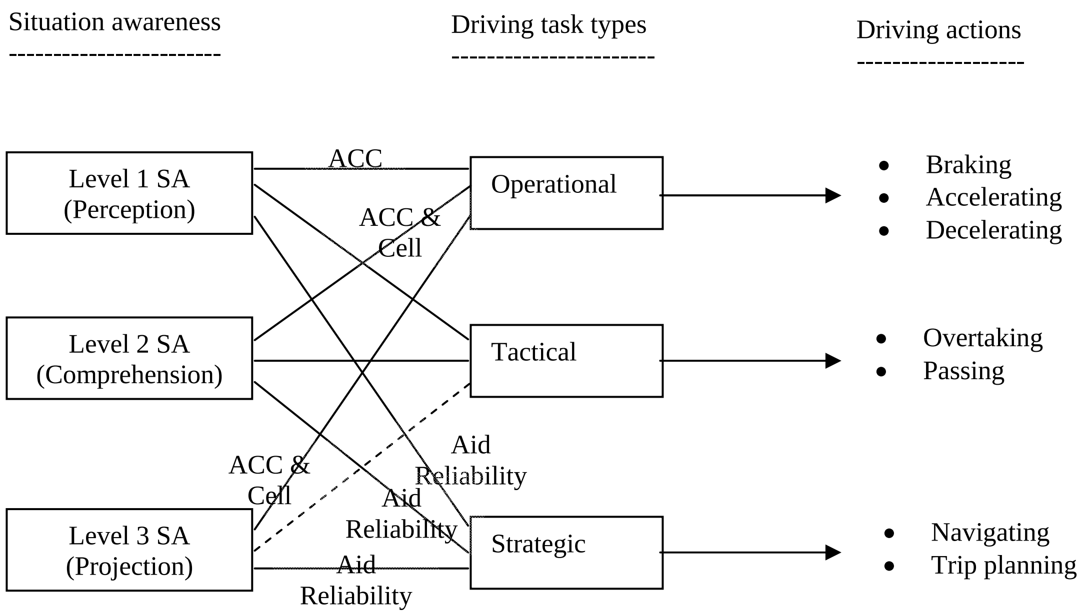


Figure 26: Updated transactional model of situation awareness and its potential influence on driving task types and driver actions

The results summarized in the updated transactional model apply to idealized driving situations (e.g., freeway driving with limited interactive traffic). Additional mental resources may be required of drivers under non-normal or hazardous driving conditions. This factor and other individual variables, like age and perceptuo-cognitive abilities, might lead to a different set of results and a more complex version of the transactional model of driver SA.

All of the correlation analyses as part of pilot study (see Ma and Kaber, 2005), and the freeway and suburb navigation driving experiments revealed positive associations between overall SA (and some specific levels of SA) with one or more dimensions of driving performance. For example, there were significant negative linear associations between total SA score and variations in headway distance and following speed in the pilot study; and a significant negative linear association between Level 3 SA scores and driving navigation errors in the navigation driving task. These additional findings demonstrate the importance of the cognitive construct to driving and further reveal the importance of the transactional model of SA in driving as a predictive tool for driving performance.

8.4 Trust measures

In Experiment B, it appeared that people had the expectation that automation would generally be more reliable than a human aid, based on their daily experiences. Consequently, participants provided higher initial trust expectations and lower ratings of expected errors for automation, as compared to human in the simulated navigation driving task. These results were in agreement with the study by Dzindolet et al. (2001), in which they found that automated aiding was perceived as more reliable than human aiding in terms of complex system operator trust. Drivers in the present study were sensitive to the different levels of navigation aid reliability. Their trust in the aids declined along with the degradations in reliability/efficiency of the information being presented. This result was consistent with Wiegman et al. (2001), and contrary to Dzindolet et al. (2001). Drivers were able to assess the efficiency of the navigation aid information by comparison with the navigation map. This helped them to calibrate their trust in the aids. When operating under imperfect navigation aiding conditions, participants perceived the human advisor as more trust worthy.

As hypothesized, when participants experienced automation inefficiency during experiment trials, their trust ratings declined more sharply than the trust ratings for the human advisor. This finding was also consistent with Wiegmann et al. (2001) and Dzindolet et al. (2001) work. The rapid drop in driver trust in the automaton aid occurred most likely because drivers expected the aid to perform at near perfect rates, leading them to pay substantial attention to errors made by the automation.

If complex system operators are not made aware of automation algorithm limitations in advance of performance, it may lead to a focus on errors, as the errors represent a violation of expectations held by the user. This eventually led to a rapid decline in the perceived reliability of, and trust in, the aid. If people know the limitations of an automation algorithm in advance, they may not lose trust in the automation so quickly (Lee and See, 2004). Human aids may be perceived as more "familiar" and this may consequently lead to decision makers having more realistic expectations of the human than machine aids, as a basis for task performance. In other words, people do not expect their human partners to be perfect. Therefore, human errors are not easily remembered and perceived; consequently, human operators of complex systems are likely to be more forgiving of incorrect of inefficient information provided by an imperfect human aid rather than by an imperfect automated aid. Lee and See (2004) suggested showing the process and algorithms of automation to operators to make the automation more understandable. This understanding could help operator build an appropriate trust and reliance in automation aids.

From an additional correlation analysis, it was revealed that a significant relation of trust ratings and driving experience occurred. The trust ratings suggested that drivers with less experience may have only been concerned with whether the navigation aid provided correct information on the location of the destination. They were less concerned with whether the navigation aid directed an optimal route to the destination. Drivers with less experience generally gave higher trust ratings for the navigation aids. However, the drivers with more driving experience were concerned with whether the navigation aid was efficient and they constantly compared the information from aids with the best possible route, according to the paper map. (An expert driver was defined as having a minimum of 10 years or 20,000 work hours experience (Johnston, 2005; Chase and Simon, 1973)).

In a further analysis, trust ratings were categorized by expert (with over 10 years driving experience) or non-expert driver (less than 10 years). Of the 10 persons who participated in the study under automation aid condition, 2 were classified as expert drivers and 8 were considered non-expert. Of another 10 persons who participated under human aid condition, 7 were classified as expert drivers and 3 were considered non-expert. Since these sample sizes were unequal, no formal statistical tests were conducted. Figure 27 presents the mean trust ratings for navigation aid type by expert and non-expert driver. Graphical analysis revealed there to be almost no difference between the two types of drivers in initial ratings and after the perfect navigation aiding trial. However, the trust ratings after the imperfect navigation aiding trials revealed that expert drivers had a sharper decline in trust ratings than non-expert. It appeared that expert drivers had a lower threshold for errors in the navigation information, and non-expert drivers had more tolerance for inefficiency in route information (a higher threshold for errors).

These observations may be related to the self-reliance or self-confidence of drivers, as Lee and Moray (1992) and Kantiowitz et al. (1997) suggested that users who have low knowledge or self-confidence in a task situation tend to trust an automated aid more. There is a need to know if driver expertise influences trust in human and/or automation aids when navigation information reliability varies.

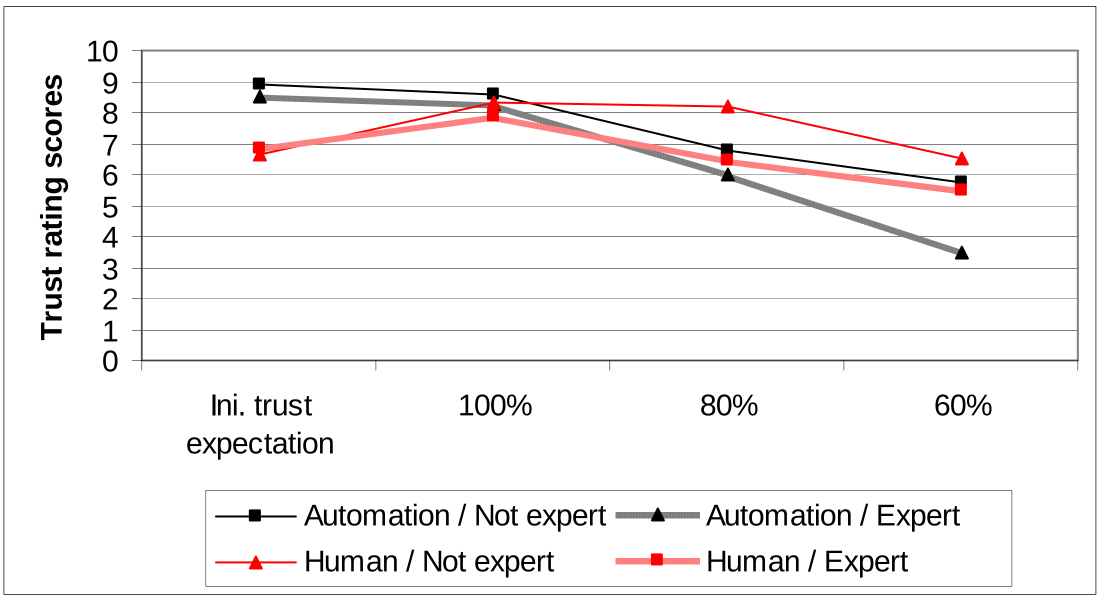


Figure 27: Mean trust ratings for navigation aid type and different driver

The correlation analysis for the navigation task revealed a positive association between the post-trial trust ratings and comprehension of the driving environment (Level 2 SA). This suggested the trust in navigation aids (or the reliability of aids) helped drivers to better comprehend the states of the driving environment more completely and accurately relative to their driving objectives.

9 Conclusions

The objectives of this study were to: (1) investigate ACC and cell phone use in driving on a direct, objective measure of SA; (2) investigate the effect of varying navigation reliability on driver SA and trust; (3) detail aspects of the new transactional model of levels of SA in driving behaviors and activities, as influenced by automation and in-vehicle device use; and (4) assess differences in driver trust in human versus automation aids in a simulated driving navigation task.

9.1 The effects of in-vehicle automation on driver performance, SA and workload

An operational definition of SA in the driving domain was developed as part of this research and it was applied as a measure of driver perception, comprehension and projection in a lead-car following task, and freeway driving and suburb navigation tasks. The results of experiments provided support for use of in-vehicle automation, including ACC and navigation aiding, under normal driving conditions for facilitating driver SA. The in-vehicle automation appeared to relieve drivers of workload in, for example, monitoring and implementing speed control and frequently checking driving directions, allowing them to develop more complete and accurate knowledge of driving environment states.

It is possible that the use of ACC or navigation aids in complex driving tasks, involving unexpected or hazardous conditions, may have different affects on driver performance and SA. The use of ACC or navigation aids under normal driving conditions, leading up to driver negotiation of a hazardous event, might cause out-of-the-control loop unfamiliarity problems that could ultimately degrade driving performance in, for example, evasive steering or high-speed braking. Drivers using ACC for some period of time could potentially be caught-off-guard by an unexpected hazard and may not be prepared to quickly deactivate the ACC and demonstrate emergency braking, etc. Future work should look at the affects of ACC and navigation aids under non-normal or critical/hazardous driving conditions.

The empirical results from the lead-car following experiment confirmed hypothesized linkages of the three levels of SA with operational driving behaviors/actions. The results on the navigation task provided support for hypothesized linkages among the three levels of SA and strategic driving behaviors/actions. All results served to detail the transactional model of SA in driving. The current transactional model can describe the affects of in-vehicle automation (e.g., ACC and navigation aids) and devices (e.g., cell phones) on driving task SA and performance. However, there remains a need to provide empirical evidence of the role of driver SA in tactical driving behavior (passing, overtaking, etc.), as part of the model. As mentioned, because of limitations in SA measurement techniques used in other research on tactical driving, additional experiments using the SAGAT method, or another objective and global measure of SA, are needed.

The results of the driver SA research presented here may lead to the definition of more comprehensive, direct objective measures of SA in driving and more accurate quantitative descriptions of the role of each level of SA in performance of the various types of driving tasks. The current study and the pilot study identified the use of in-vehicle automation and its reliability as underlying factors in the linkages of the transactional SA model. Of course, there may be many other individual, technological and driving system factors involved in the model linkages. Further validation of the transactional model and identification of other underlying factors may lead to the model being used for predictive purposes in future systems design. For example, it could be used to predict SA levels in driving with new forms of in-vehicle automation sharing characteristics with current forms of automation.

Related to the findings on in-vehicle device use and SA, legislation banning the use of cell phones in driving tasks continues to increase as a result of more and more accidents attributed to cell phone use. At least 45 States in the U.S. have proposed bills concerning the use of cellular phones in automobiles since 1995 (Sundeen, 2001). According to the Royal Society for the Prevention of Accidents (RoSPA) (2005), some parts of the U.S. have passed laws to ban cell phone use while driving. For example, the first law regulating cellular telephone use in driving was in Brooklyn, Ohio in March of 1999. Following this, two laws were passed in New Jersey and New York. At least 13 municipalities in New York State have used city ordinances to ban the use of mobile phones while driving within city limits. Drivers are prohibited from talking on hand-held mobile phones while operating a motor vehicle. Fines for violation of these laws range up to $100 in New York State. According to Strayer (2005), the greatest problem caused by cell phone use while driving is mental distraction. Strayer (2005) states, “if you are driving and on a cell phone, you are about four times more likely to be in an accident, similar odds to those driving with a blood alcohol level of 0.08”. As demonstrated in this research mental distraction in driving, as a result of in-vehicle automation or cell phone use can lead to significantly decrements in driver SA that are linked to specific driving behaviors and performance.

Potentially compounding the danger associated with in-vehicle cell phone use, recently major cell phone companies have begun to contract with television (TV) companies to deliver TV programs to cell phone users. For example, *Qualcomm* is working to build a system to transmit live TV and stored clips directly to handsets, independent of cellular networks; *RealNetworks* is entering into a partnership with Cingular to deliver high-quality video with interactive capabilities for cell phone delivery; and *MobiTV* is releasing MobiTV2, a new platform providing on-demand premium videos for mobile viewers (e.g., using cell phones) (Downs, 2005). The use of this type of technology/information in the context of driving may create greater distraction for drivers and greater danger, compared to only talking on a cell phone while following other cars, navigating new roads, etc. The mental resources of drivers may be stretched even further in the future by such mobile device capabilities. This will likely instigate more debates and laws against cell phone use during driving. There is a need to continue to investigate how drivers achieve and maintain SA in the presence of cell phone use and in-vehicle automaton as a basis for driving performance. The resulting knowledge should be applied to the development of future technologies, or the redesign of existing devices, and reflected in any state and/or federal regulations on in-vehicle device use.

9.2 The effects of in-vehicle automation on driver trust

This study revealed differences in human trust in in-vehicle automation and human advisors for navigation in driving. Automation was generally expected to be more reliable and make fewer errors than a human in the simulated navigation driving task. However, when participants experienced automation errors or inefficiency in route planning, their trust in automation declined more sharply than trust in the human advisor. It also appeared that the number of years a person had been driving influenced the degree of trust in navigation aids. More experienced drivers tended to be less trusting of unreliable automation.

Finally, the study also demonstrated that drivers had better comprehension of the driving environment if they had higher trust in navigation assistance. Imperfect navigation aiding reduced driver trust and degraded driver SA and, consequently, strategic driving performance.

9.3 Caveats

There are some limitations of this study that should be noted with respect to using the results as a basis for designing or making decision the use of in-vehicle automation. The first limitation was the use of a medium fidelity driving simulation for all experiments. The simulator was a fixed-base setup providing no kinesthetic motion and there was no interactive traffic represented in the simulation. Drivers may behave differently in actual operational settings because of the serious consequences of having an accident. The second limitation of this study was related to the scheduling of SAGAT freezes in the freeway driving and navigation task experiments. Participants knew simulation freezes and queries would occur roughly at the end of the freeway or navigation driving tasks, and they may have taken advantage of this knowledge to prepare for SA questions. Another limitation of the driving navigation task experiment was the fixed order of presentation of the navigation aid reliability conditions. Driver trust in navigation aids was investigated by systematically degrading the reliability of navigation information across trials. This was necessary to study changes in driver trust across the conditions, but the trial order may have influenced the measurement of other dependent variables, including driver SA.

9.4 Future research directions

On the basis of this study, directions of future research include investigating SA in a tactical driving task and further describing the relationships among the various levels of SA, tactical driving behavior and relevant actions/performance in the transactional model of SA for driving. There may be many individual, technological and driving system factors involved in the linkages in the transactional model, extending beyond the use of ACC systems and automated navigation aids. There is a need to use highly realistic and complex driving tasks to investigate such linkages in the model. The current research used a typical, abstract simulation of two basic driving behaviors (i.e., operational and strategic behaviors). There is a need for further future research that investigates SA in combinations of driving behaviors in high-fidelity simulators.

Future research aimed at incrementally advancing the present study includes introducing additional response measures, for example, eye tracking. Participants frequently visualized the driving interface, the navigation information from the additional computer display, and the hard copy of the suburb map during experiment trials. It appeared that drivers adopted different visual scanning strategies to balance performance in multitasking (i.e., driving and navigating). Eye tracking data could provide more evidence on the relationships among SA, workload and driving performance by detailing what drivers attend to, when. Eye tracking measures would allow for more insight into the transactional model of SA in driving.

Another direction of future research would involve developing a high-fidelity driving simulation and running an experiment with a random ordering of navigation aid reliability conditions to investigate the affects of automated or human aiding on driver trust, SA, workload and performance. A high-fidelity simulation with more interactive functions, both internal (e.g., ACC set and cancel) and external (e.g., traffic light signal, etc.) would be helpful in terms of promoting the generalizability of results. The random ordering of automation reliability conditions might serve to provide clear statistical conclusions on the effects of navigation aiding reliability on driver SA and performance. Another experiment could be developed to address or assess the affect of trial order on the response measures. A study could be designed requiring certain participants to perform under a constant reliability condition and for comparison of results to be made with a treatment group, which was exposed to all aid reliability levels. The experience of drivers could also be formally considered as an independent variable in such research, particularly for examining the implications on trust in in-vehicle automation under the various navigation aid reliabilities.

10 References

Billings, C.E. (1991). *Human-centered aircraft automation: A concept and guidelines* (NASA Technical Memorandum 103885). Moffet Field, CA: NASA-Ames Research Center.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ: Lawrence Erlbaum Association.

Chase, W. and Simon, H. (1973). Perception in Chess. *Cognitive Psychology, 4*, 55-81.

Chen, H. J. and Lin, C. J. (2003). Effects of wireless communication on driving performance using a desktop driving simulator. In *Proceedings of XVth Triennial Congress of the International Ergonomics Association Conference (CD-ROM)*, August 24-29, 2003, Seoul, Korea.

Corriotore, C.L., Kracher, B. and Wideenbeck, S. (2003). Online trust: Concept, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*, 737-758.

Downs, S. (2005). Cell phones to deliver television programming. Retrieved November 29, 2005, from http://myst-technology.com/mysmartchannels/public/item/104373

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G. and Beck, H.P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58 (6),* 697-718.

Dzindolet, M.T., Pierce, L.G., Beck, H.P. and Dawe, L.A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors, 44 (1)*, 79-94.

Dzindolet, M.T., Pierce, L.G., Beck, H.P., Dawe, L.A. and Anderson, B.W. (2001). Predicting misuse and disuse of combat identification systems. *Military Psychology, 13 (3),* 147-164.

Edwards, E. (1976). Some aspects of automation in civil transport aircraft. In T.B. Sheridan & G. Johannsen (Eds.), *Monitoring behavior and supervisory control.* New York: Plenum.

Edwards, M. (2001). Driver distraction and safety. In: *Implications for Telematic Devices*. AAA White Paper, Lake Mary, FL.

Endsley, M.R. (1993). Situation awareness and workload: Flip sides of the same coin. In R.S. Jensen and D. Neumeister (Eds.), *Porceedings of the Seventh International Symposium on Aviation Psychology* (pp. 906-911). Columbus, OH: Department of Aviation, The Ohio State University.

Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37(1),* 32-64.

Endsley, M.R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors, 37(1),* 65-84.

Endsley, M. R., and Jones, D. G. (1995). *Situation awareness requirements analysis for TRACON air traffic control* (Tech. Report TTU-IE-95-01). Atlantic City, NJ: Technology Center, Federal Aviation Administration.

Endsley, M. R. and Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automaiton. *Human Factors, 37,* 381-394.

Gugerty, L.J. and Tirre, W. C. (2000). Individual difference in situation awareness. *In Situation Awareness Analysis and Measurement (Edited by Endsley, M.R. and Garland, D.J.).* pp 249-276.

Gugerty, L., Rando, C., Rakauskas, M., Brooks, J., and Olson, H. (2003). Differences in remote versus in-person communications while performing a driving task. In *Proceedings of XVth Triennial Congress of the International Ergonomics Association Conference (CD-ROM)*, August 24-29, 2003, Seoul, Korea.

Hancock, P.A., Simmons, L., Hashemi, Howarth and Ranney (1999). The effects of in-vehicle distraction on driver response during a critical driving maneuver. *Transportation Human Factors, 1(4),* 295-309.

Hancock, P.A., Lesch, M. and Simmons, L. (2002). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention.* 1-14.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). North-Holland Elsevier Science: Amsterdam.

Hwang, P. and Burgers, W.P. (1997). Properties of trust: An analytical view. *Organizational Behavior and Human Decision Process, 69(1),* 67-73.

Itoh, M., Inagaki, M., Shiraishi, Y., Wateanabe, T. and Takae, Y., (2005). Contributing factors for mode awareness of a vehicle with a low-speed rang and a high-speed range ACC systems. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Jerome, C.J., Ganey, H.C., Mouloua, M., and Hancock, P.A.(2002). Driver workload response to in-vehicle device operations. *International Journal of Occupational Safety and Ergonomics, 8*(44). 539-548.

Johnston, R. (2005). Integrating methodologists into teams of substantive experts. Retrieved November 9, 2005, from http://www.cia.gov/csi/studies/vol47no1/article06.html#fn2

Kantowitz, B.H., Hanowski,R.J. and Kantowitz,S.C. (1997). Driver reliability requirements for traffic advisory information. In:Noy,Y.I.(Ed.), *Ergonomics and Safety of Intelligent Driver Interfaces*. Lawrence Erlbaum Associates, Mahwah,NJ,pp.1-22.

Lee, J.D. and Moray, N. (1992). Trust, control strategies and allocation of function in human machine systems. *Ergonomics, 22(6),* 671-691.

Lee, J. D., and See, K. A. (2004). Trust in Technology: Designing for Appropriate Reliance. *Human Factors, 46(1),* 50-80.

Lewicki, R.J. and Bunker, B.B (1995). Trust in relationships: A model of development and decline. In B.B. Bunker and J. Z. Rubin (eds.), *Conflict, Cooperation, and Justice: Essays Inspired by the Works of Morton Deutsch,* 133-173. San Francisco, CA: Jossey-Bass.

Lewis, D.J. and Weigert, A. (1985). Trust as a social reality. *Social Forces, 63,* 967-985.

Ma, R. and Kaber, D.B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomic, 35(10),* 939-953.

Ma, R., Kaber, D.B. and Chow, M.Y. (2004). A “distance”-based concept of automation for human-robot interaction. In *Proceedings of 2nd HPSAA Conference (CD-ROM),* March 22-25, 2004, Daytona Beach, FL.

Maurel, D. and Donikian, S. (2001). ACC systems-overview and examples. *Intelligent Vehicle Technologies (Edited Vlacic, L., Parent, M. and Harashima, F.).* pp 423-441.

Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., and Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Human Factors, 40(1),* 136-149.

Matthews, M.L., Bryant, D.J., Webb, R.D. and Harbluk, J.L. (2001). Model for situation awareness and driving. *Transportation Research Record, 1779,* 26-32.

Mayer, R.C., Davis,J.H., and Schoorman,F.D. (1995).An integrative model of organizational trust. *Academy of Management Review 20 (3),* 709-734.

Muir, B.M. (1994). Trust in automation: Part 1, Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics, 37(11),* 1905-1922.

Muir, B.M. and Moray, N. (1996). Trust in automation: Part 2, Experimental studies of trust and human intervention in a process control simulation. *Ergonomics, 39,* 429-460.

Navon, D. and Gopher, D.,(1979). On the economy of the human information processing system. *Psychological Review, 86*, 214-255.

Parasuraman, R. and Riley, V. (1997). Human and automation: Use, misuse, disuse, abuse. *Human Factors, 39(2),* 230-253.

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

Parker, H.A., Malisia, A.R, and Rudin-Brown, C.M. (2003). Adaptive cruise control and driver workload. In *Proceedings of XVth Triennial Congress of the International Ergonomics Association Conference (CD-ROM),* August 24-29, 2003, Seoul, Korea.

Parsons, H.M. (1985). Automation and the individual: Comprehensive and comparative views. *Human Factors, 27,* 99-111.

Rasmussen, R. (1983). Skills, rules and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, *Man, and Cybernetics, 13(3), 257-266.*

RoSPA (2005). The risk of using a mobile phone while driving. Retrieved November 29, 2005, from http://www.rospa.com/roadsafety/info/mobile_phone_report.pdf

Rudin-Brown, C.M., Parker, H.A, Malisia, A. R., (2003). Behavioral adaptation to adaptive cruise control. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Sheridan, T. B. (2002). *Human and Automation: System Design and Research Issues*. A John Wiley & Sons.

Shinar, D., and Ben-Shoham, I. (1998). How automatic is manual gear shifting. *Human Factors, 40(40),* 647-654.

Stanton, N.A., and Marsden, P (1996). From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. *Safety Science*, *24*, 35-49.

Stevens, W.B. (1997). Evolution to an automated highway system. *In Automated Highway Systems (Edited by Ioannou, P.A.).* pp 109-124.

Strayer, D.L. (2005). Drivers using mobile phones 4 times more likely to have accidents. *The Ergonomist*, November 2005.

Sukthankar R.(1997). *Situation Awareness for Driving in Traffic*. PhD thesis, Carnegie Mellon University, January 1997.

Sundeen, M. (2001). Cell Phone and Highway Safety: 2000 State Legislature Update. *National Conference of State Legislatures*, May 2001, UK

Taylor, R.M., Selcon, S.J. and Swinden, A.D. (1995). Measurement of situational awareness and performance. In *Human Factors in Aviation Operations,* R. Fuller, N. Johnston and N. McDonald (Eds.) (Aldershot: Avebury).

Waern, Y. and Ramberg, R. (1996). People's perception of human and computer advice. *Computer in Human Behavior, 12(1),* 17-27.

Ward, N.J. (2000). Automation of task processed: an example of intelligent transportation systems. *Human Factors and Ergonomics in Manufacturing, 10(4),* 395-408.

Wickens, C.D. (1984). Processing resources in attention. In *Varieties of Attention.* R. Parasuraman and D. Davis, Eds. New York: academic.

Wickens, C.D. and Hollands, J.G. (2000). *Engineering Psychology and Human Performance, 3rd edn.*. Upper Saddle River, NJ: Prentice Hall.

Wiegmann, D.A., Rich, A. and Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on user's trust and reliance. *Theoretical Issues in Ergonomics Science, 2(4)*, 352-367.

Wiener, E.L. (1985). Cockpit automation: In need of a philosophy. In *Proceedings of the 1985 Behavioral Engineering Conference (pp. 369-375).* Warrendale, PA: Society of Automotive Engineers.

Wiener, E.L. (1988). Cockpit automation. In E.L. Wiener, & D.C. Nagel (Eds.). *Human Factors in Aviation* (p. 433-461). San Diego, CA: Academic.

Wiener, E.L. (1993). Life in the second decade of the glass cockpit. In R.S. Jensen and D. Neumeister (Eds.), *Porceedings of the Seventh International Symposium on Aviation Psychology* (pp. 1-11). Columbus, OH: Department of Aviation, The Ohio State University.

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

Appendix A: Example SA Questionnaire in Pilot Study

SITUATION AWARENESS QUESTIONNAIRE

Situation awareness and workload in driving while using adaptive cruise control and a cell phone

1. What is the color of the vehicle in the right lane behind your car? (L1)
   No car
   Grey
   White
   I don't know

2. What was the last road sign you saw? (L1)
   Pedestrian crossing
   Deer crossing
   Railroad
   Speed limit
   Slow sign
   I don't know

3. What is your vehicle position? (L1)
   In left lane
   In right lane
   Between left and right lanes
   I don't know

4. How long has it been since you passed the last sign? (L2)
   2-7 seconds
   8-12 seconds
   13-17 seconds
   Bigger than 17 seconds

5. At this moment, do you need to accelerate to catch the leading car? (L2)
   Yes
   No
   I don't know

6. How much do you need to accelerate to catch the leading car? (L2)
   None
   Slow
   Moderate
   Fast

7. When should you accelerate to catch the leading car? (L3)
   I don't need to accelerate
   Immediately
   In the near future
   Later

8. How long do you need to accelerate to catch the leading car? (L3)
   None
   1-4 seconds
   5-8 seconds
   9-12 seconds
   13-16 seconds

9. When should you decelerate to back-off the leading car? (L3)
   I don't need to decelerate
   Immediately
   In the near future
   Later

Do not write below this line. Experimenter use only.

Subject Number: __________
Trial Number: __________
Freeze Number: ____1______

Appendix B: Driving Simulator Goal-Directed-Task-Analysis

A GDTA is accomplished by first identifying major goals required for accomplishing the task. Secondly, major subgoals are determined that are essential for meeting the goals. Then, major decisions that are associated with each subgoal are developed and used to determine SA requirements for accomplishing the task. These requirements focus not only on what data the driver/operator needs, but also on how that information is integrated or combined to address each decision associated with the three levels of situation awareness (SA).

The following GDTA describes the goals and information requirements required to accomplish the proposed driving task; using the methods described by Endsley and Jones (1995). Each major decision and SA requirement associated with a subgoal represents a level of information processing (perception, comprehension, or projection), and they were used to develop SAGAT queries. This analysis assumes there are no operator errors.

Directly following the GDTA is a description of a plan for accomplishing the goals within the driving simulator.

(Note: Some decisions or questions make reference to other information/SA requirements. These requirements are emboldened.)

Goal	Subgoal	Decisions/SA requirements	Level of SA

0. follow lead car safely in a driving simulator

1. perceive driving environment

1.1 observe lead car behavior

what is the position of the car in front of me?	Level 1
user car position	
distance from user car to lead car	
what is the speed of the car in front of me?	Level 1
user car position	
lead car position	
distance change	
is the car in front of me making lane change?	Level 1
user car position	
lead car position	

1.2 perceive traffic

is there any car behind user car?	Level 1
is there any car on the left/right of user car?	Level 1
user car position	
view of the rear, left and right mirrors	
amount of cars, color of cars	
What is the status of cars behind, on left/right of user car?	Level 1
User car position	
Other cars position	
Gaining	
Losing	
Lane maintenance	

1.3 perceive signals & signage

what is the nearest traffic sign to user car?	Level 1
what was the last road sign you saw?	Level 1
what is the distance to sign/signal?	Level 1
what does sign say?	Level 1
what is the color of sign?	Level 1
user car position	
sign position	
is there a pedestrian sign in view?	Level 1
How far is sign from car?	Level 2
when will your vehicle arrive at Pedestrian crossing sign?	Level 3
user car position	
sign position	
user car speed	
is there a deer sign in view?	Level 1
How far is the sign from car?	Level 1
when will your vehicle arrive at the Deer sign?	Level 3
user car position	
sign position	

user car speed

…..(more similar critical decisions)

2. establish appropriate position relative to lead car

what is distance from lead car? Level 2
lead car position (1.1)
user car position
what is the heading of lead car? Level 2
lead car position (1.1)
user car position

3. select driving action/maneuver
3.1 speed control (only for No-ACC control)
3.1.1 accelerate to catch leading car
do you need accelerate to catch leading car? Level 2
how much do you need to accelerate? Level 2
when do you need to accelerate? Level 3
how long do you need to accelerate? Level 3
lead car position (1.1)
lead car speed (1.1)
user car position
speed limit (1.3)

3.1.2 decelerate to back-off lead vehicle
do you need decelerate to achieve comfortable following distance? Level 2
how much do you need to decelerate ? Level 2
when do you need to decelerate? Level 3
how long do you need to decelerate? Level 3
lead car position (1.1)
lead car speed (1.1)
user car position
speed limit (1.3)

3.2 orientation control
3.2.1 change lanes to follow lead car
do you need change lane to follow the lead car? Level 2
are you changing lane? Level 1
lead car position (1.1)
user car position
lead car lane change (1.1)

3.2.2 turn left for curve
when will you need to turn left for curve? Level 2
user car position
curve position
how much do you need to turn? Level 3
how long do you need to turn? Level 3

lead car position (1.1)
user car position
curve position

3.2.3 turn right for curve

when will you need to turn right for curve? — Level 2
- user car position
- curve position

how much do you need to turn wheel? — Level 3

how long do you need to turn wheel? — Level 3
- lead car position (1.1)
- user car position
- curve position

4. implement driving action

4.1 speed control (only for No-ACC control)

4.1.1 apply gas

how much do you need to press gas pedal? — Level 2
- lead car position (1.1)
- user car position
- distance between user car and lead car

4.1.2 apply brake

how much do you need to press brake? — Level 2
- lead car position (1.1)
- user car position
- distance between user car and lead car

4.2 orientation control

4.2.1 turn wheel left

how much do you need to turn wheel? — Level 2
- lead car position (1.1)
- lead car heading
- user car position

4.2.2 turn wheel right

how much do you need to turn wheel? — Level 2
- lead car position (1.1)
- lead car heading
- user car position

4.3 monitor state of vehicle/perceive system feedback

is the user car reducing speed? — Level 1
- lead car position (1.1)
- user car position

is the user car reducing speed enough? — Level 2
- lead car position (1.1)
- user car position

distance between user car and lead car

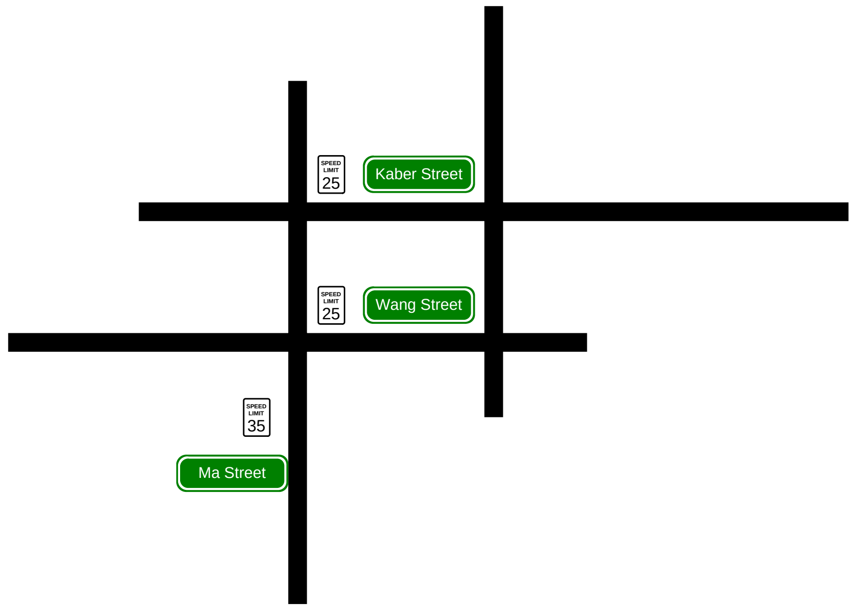
is the user car changing orientation? — Level 1
user car position
lane position
is the user car changing orientation enough? — Level 2
user car position
lane position
lead car position

what is the car position (in lane/on road)? — Level 1
user car position
lane position
is your car gaining on lead car? — Level 2
user car position
lead car position
what is the status of your car? — Level 2
(following, passing, stopping, turning, etc.)
user car position
lane position

Plan 1: do 1.1 – 1.2 – 1.3 – 1.4, maybe not always strictly repeat the sequence

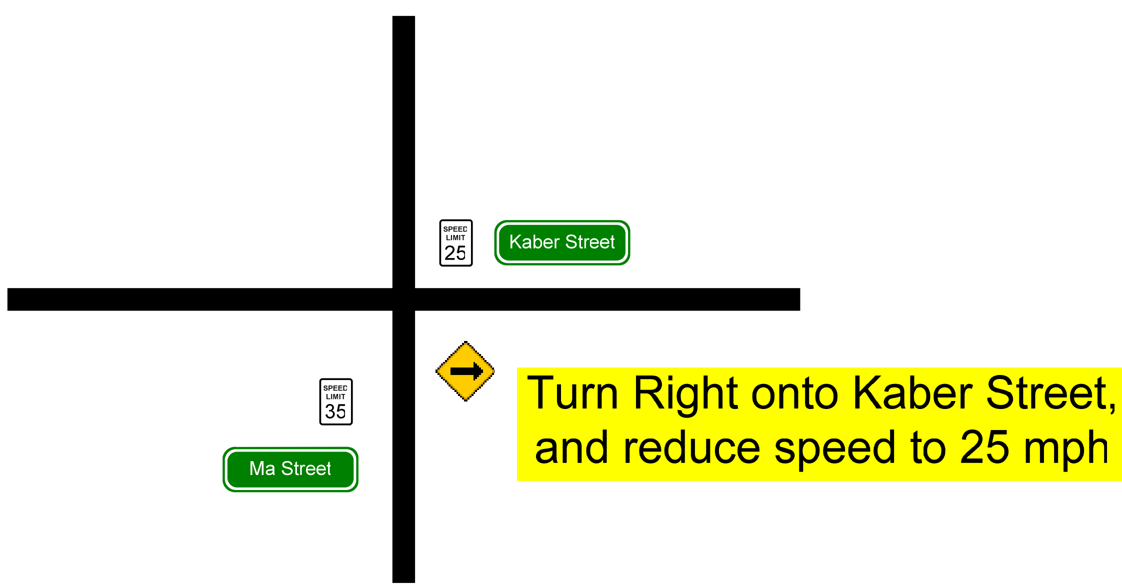
Appendix C: Script and Display Examples of Human and Automation Navigation Aids

Map and relative script for Human Aid



"Hello driver"
"I am calling to instruct you in driving directions in your navigation task."
"First, go straight and pass Wang Street."
"Now, turn RIGHT onto Kaber Street. Your diving speed should be reduced to 25 mph."
……

Display for Automation Aid



Appendix D: Telemarketing Script Example

Administered to Participants during Telemarketing Survey

Hello, driver, I am calling on behalf of a retail store in your area. We are conducting a short survey on shopping. During the course of your drive, I would like to ask you a few short questions. Are you ready?

- ☐ No - (repeat the above)
- ☐ Yes - (go to question 1)

[Pause for 1 to 1.5 minutes.]

1. How often do you go shopping at malls?

2. When was the last time you went to a shopping mall?

3. What did you buy during your last shopping spree?

Thank you for your responses. I may call you back in a few minutes.

[Pause for 1 to 1.5 minutes.]

Hello again driver. I would like to ask you a few more questions.

4. What is your favorite shopping center/mall?

5. How much do you normally spend when you go shopping?

6. When do you normally go shopping, weekdays or weekends?

Thank you for your responses. I may call you back in a few minutes.

[Pause for 1 to 1.5 minutes.]

Hello again driver. I would like to ask you a few more questions.

7. Do you go shopping with your friends or by yourself?

8. Do you drive a car or take public transportation to go shopping?

9. What is the closest shopping center to your apartment or house?

Thank you for your responses. I may call you back in a few minutes.

[Pause for 1 to 1.5 minutes.]

Hello again driver. I would like to ask you a few more questions.

10. What is your primary purpose when you go shopping, buying merchandise, browsing, seeing friends?

11. Do you frequently go to bookstores when you go shopping in a shopping center?

12. How much do you normally spend in bookstores per visit?

Thank you for your responses. I may call you back in a few minutes.

[Pause for 1 to 1.5 minutes.]

Hello again driver. I would like to ask you a few more questions.

13. What is your favorite book?

14. How much time do you spend reading your favorite book?

15. Can you give me the name of some of your favorite books?

Do not write below this line. Experimenter use only.

Subject # _________

Appendix E: Example SA Questionnaire in Navigation Driving Task

SITUATION AWARENESS QUESTIONNAIRE
THE EFFECT OF IN-VEHICLE AUTOMATION AND RELIABILITY ON DRIVER SITUATION AWARENESS AND TRUST

1. What was the last road sign you saw? (L1)
   - Pedestrian crossing
   - Deer crossing
   - Railroad
   - Speed limit
   - Slow sign
   - Stop sign
   - I don’t know

2. What was the color of the vehicle directly behind your car at the time the simulation stopped? (L1)
   - No car
   - Grey
   - White
   - I don’t know

3. What was your vehicle speed (mph) at the time the simulation stopped? (L1)
   - Less than 25
   - 25-30
   - 30-35
   - 35-40
   - 40-45
   - 45-50
   - More than 50

4. How long has it been since you passed the last turn in navigating the city suburb? (L2)
   - Less than 30 seconds
   - 30 seconds - 1 minute
   - 1 - 1.5 minutes
   - 1.5 - 2 minutes
   - 2 – 2.5minutes
   - 2.5 - 3 minutes
   - More than 3 minutes

5. How long has it been since you passed the last road sign? (L2)
   - 1-5 seconds
   - 5-10 seconds
   - 10-15 seconds
   - 15-20 seconds
   - 20-25 seconds
   - More than 25 seconds

6. In which direction from your vehicle was your destination (building) located when you passed the last turn in the city suburb? (L2)
   On the left
   On the right
   Right in front of me
   Behind me
   I don't know

7. When will your vehicle reach the next road sign in the simulation? (L3)
   No sign in sight
   1-5 seconds
   5-10 seconds
   10-15 seconds
   15-20 seconds
   More than 20 seconds

8. What was the optimal navigation route to reach your destination when you passed the intersection of Kaber St. and Ma St. (what route would have generated the shortest drive time)? (L3)
   Go to Ma St., then Riley Rd, then McDowell St
   Go to Kaber St., then Noa Dr., then Riley Rd., then McDowell St.
   Go to Ma St., then Avent Ferry St.
   I don't know

9. How much longer until you finish driving through the suburban area? (L3)
   Less than 1 minute
   1-2 minutes
   2-3 minutes
   3-4 minutes
   4 -5 minutes
   More than 5 minutes

Do not write below this line. Experimenter use only.

Subject Number: __________
Trial Number: ____2______
Experiment: _____B_____

Appendix F: NASA-TLX Workload Assessment Forms

NASA-TLX Workload Factor Definitions

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

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

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

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

Frustration
How insecure, discouraged, irritated, and annoyed versus secure, gratified, content and complacent did you feel during the task?

Effort
How hard did you have to work (mentally and physically) to accomplish your level of performance?

Subjective Comparison of Demand Factors: NASA-TLX Survey

Indicate the demand of greater importance by circling its label on each line directly below.

Mental Demand / Physical Demand

Mental Demand / Temporal Demand

Mental Demand / Performance

Mental Demand / Effort

Mental Demand / Frustration

Physical Demand / Temporal Demand

Physical Demand / Performance

Physical Demand / Effort

Physical Demand / Frustration

Temporal Demand / Performance

Temporal Demand / Frustration

Temporal Demand / Effort

Performance / Frustration

Performance / Effort

Frustration / Effort

Do not write below this line. Experimenter use only.

Subject # _________ Trial # _________

Subjective Rating of Perceived Workload: NASA-TLX Survey

The effect of in-vehicle automation and reliability on driver situation awareness and trust

---

Indicate the level of demand experienced during the navigation task for each of these factors by drawing a straight vertical line on the scale directly below.

Mental Demand

Low High

Physical Demand

Low High

Temporal Demand

Low High

Performance

Low High

Frustration

Low High

Effort

Low High

Do not write below this line. Experimenter use only.

Subject # __________ Trial # __________

Appendix G: Initial Trust Questionnaire when Using Human Aid

Administered to Participants in the Initial Expectation Condition

1. How well do you think the human aid will perform during the trials?

Poor				Moderate					Excellent
1	2	3	4	5	6	7	8	9	10

2. How many errors do you think the human aid will make during the trials?

The human aid will make about ______________ errors.

Do not write below this line. Experimenter use only.

Subject # _________

Appendix H: Initial Trust Questionnaire when Using Automation Aid

<u>Administered to Participants in the Initial Expectation Condition</u>

1. How well do you think the automation aid will perform during the trials?

Poor				Moderate					Excellent
1	2	3	4	5	6	7	8	9	10

2. How many errors do you think the automation aid will make during the trials?

<u>The automation aid will make about ______________ errors.</u>

Do not write below this line. Experimenter use only.

Subject # _________;

Appendix I: Trust Questionnaire when Using Human Aid

<u>Administered to Participants after Each Trial</u>

1. To what extent do you believe you can trust the decisions of the human aid?

Not at all					Some				Completely
1	2	3	4	5	6	7	8	9	10

2. How would you rate the performance of the human aid relative to your performance?

Poor					Moderate				Excellent
1	2	3	4	5	6	7	8	9	10

3. Please explain your responses to Items # 1 and 2. Do not worry about spelling or grammatical errors--just let us know why you think the aid performed at the rated level.

   Explain to # 1:

____
____
____
____
____
____

   Explain to # 2:

____
____
____
____
____
____

Do not write below this line. Experimenter use only.

Subject # __________ Trial # __________

Appendix J: Trust Questionnaire when Using Automation Aid

Administered to Participants after Each Trial

1. To what extent do you believe you can trust the decisions of the automation aid?

Not at all					Some				Completely
1	2	3	4	5	6	7	8	9	10

2. How would you rate the performance of the automation aid relative to your performance?

Poor					Moderate				Excellent
1	2	3	4	5	6	7	8	9	10

3. Please explain your responses to Items # 1 and 2. Do not worry about spelling or grammatical errors--just let us know why you think the aid performed at the rated level.

Explain to # 1:

____
____
____
____
____
____

Explain to # 2:

____
____
____
____
____
____

Do not write below this line. Experimenter use only.

Subject # __________ Trial # __________

Appendix K: Informed Consent Form

NORTH CAROLINA STATE UNIVERSITY
INFORMED CONSENT FORM for RESEARCH

Title of Study The effect of in-vehicle automation and reliability on driver situation awareness and trust
Principal Investigator Ruiqi Ma Faculty Sponsor (if applicable) David B. Kaber
----

We are asking you to participate in a research study. The purpose of this study is to: (1) investigate the implications of ACC and cellular phone use in driving on a direct and objective measure of SA; (2) investigate the effect of varying reliability of in-vehicle automation on driver SA and trust; and (3) to assess differences in human trust in a human aid versus an automation aid in a simulated driving task.

INFORMATION
If you agree to participate in this study, you will be asked to follow these procedures as part of your participation: (1) instruction in a virtual reality based driving simulation (15 min.); (2) training under a no-ACC driving condition without cellular use (10 min.); (3) familiarization with a situation awareness questionnaire and workload questionnaire (10 min.); (4) a short break (5 min.); (5) an experimental testing session (25 min.); (6) a second experimental testing session (25 min.); (7) a third experimental testing session (25 min.); and (8) a fourth experimental testing session (25 min.). The experiment will require approximately 3.5 hours of your time.

RISKS
The risks to subjects associated with participation in this study are unlikely and minimal. They include: (1) possible soreness of the hand and leg muscles from extensive use of a steering wheel interface and accelerator and brake pedals; and (2) potential visual strain and/or fatigue in viewing the simulation display through conventional monitors for extended periods. These risks are not substantially different from those associated with everyday PC use 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 conditons persist, your participation in the experiment will be terminated.

BENEFITS
There are no direct benefits of the research. You may derive some indirect benefit including an understanding of human factors research methods and insight into the general effects of in-vehicle automation and cellular phone use on driving. You will receive an ergonomics lab t-shirt as compensation for your participation at the close of the study. You may also receive a gift certificate of $50 if you achieve the highest level of performance in the experiement tasks, as compared to all other students.

CONFIDENTIALITY
The information in the study records will be kept strictly confidential. Data will be stored securely and will be made available only to persons conducting the study. No reference will be made in oral or written reports that could link you to the study.

CONTACT
If you have questions at any time about the study or the procedures, you may contact the faculty sponsor of the study, Dr. David Kaber at the Department of Industrial Engineering, Box 7906, NCSU, or 919-515-3086**. If you feel you not been treated according to the information in this form, or your**

rights as a participant in research have been violated during the course of this project, you may contact Dr. Matthew Zingraff, Chair of the NCSU IRB for the Use of Human Subjects in Research Committee, Box 7514, NCSU Campus (919/513-1834) or Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus (919/513-2148).

PARTICIPATION
Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed at your request.

CONSENT
"I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time."

Subject's signature____ Date _________________

Investigator's signature___________________________________ Date _________________

SUBJECT SURVEY

The effect of in-vehicle automation and reliability on driver situation awareness and trust

The intended purpose of this form is to establish a subject profile based on volunteered anthropometric data. Please complete the sheet to the best of your knowledge following the example formats indicated in the parentheses adjacent to each data field label.

Name: __________________________; Age (XX yr.): _____________; Gender (M/F): __________ ;

Email address: ____;
(For contact purposes, in the event you win the gift certificate.)

Corrected Visual Acuity: Left Eye (XX/XX): ___________
Right Eye (XX/XX): ____________

Video Game Experience: 1 2 3 4 5
None Occasional Frequent

PC Experience: 1 2 3 4 5
None Occasional Frequent

VR Experience: 1 2 3 4 5
None Occasional Frequent

Driving Experience: 1 2 3 4 5
None Occasional Frequent

Cruising Use while Driving: 1 2 3 4 5
None Occasional Frequent

Cell Phone Use while Driving: 1 2 3 4 5
None Occasional Frequent

How often do you use map while Driving: 1 2 3 4 5
None Occasional Frequent

Hong long have you driven? _________

Do you own vehicles with telematics (i.e, adaptive speed control, direction control, or navigation systems)?

Yes / No (circle the right one)

DO NOT WRITE BELOW THIS LINE.

Subject #: ______