Fall 2010

Context conditions drivers' disposition towards alarms

Monica Lees
University of Iowa

Recommended Citation

Lees, Monica. "Context conditions drivers' disposition towards alarms." PhD (Doctor of Philosophy) thesis, University of Iowa, 2010.
https://doi.org/10.17077/etd.188inyqs

Follow this and additional works at: https://ir.uiowa.edu/etd
Part of the Industrial Engineering Commons

CONTEXT CONDITIONS DRIVERS' DISPOSITION TOWARDS ALARMS

by

Monica Lees

An Abstract

Of a thesis submitted in partial fulfilment
of the requirements for the Doctor of
Philosophy degree in Industrial Engineering
in the Graduate College of
The University of Iowa

December 2010

Thesis Supervisor: Professor John D. Lee

ABSTRACT

Collision warning systems represent a promising means to reduce rear-end crash involvement. However, these systems experience failures in the real-world that may promote driver distrust and diminish drivers' willingness to comply with warnings. Recent research suggests that not all false alarms (FAs) are detrimental to drivers. However, very few studies have examined how different alarms influence different driving populations.

The purpose of this research was to examine how younger, middle-aged, and older drivers (with and without UFOV impairments) evaluated and responded to four different alarm contexts – false alarm (FA), nuisance alarm (NA), unnecessary alarm (UA) and true alarm (TA) – when they did and did not receive warnings. FA contexts represent out-of-path conflict scenarios where it is difficult for the driver to identify the source of the alarm. NA contexts represent out-of-path conflict scenarios that occur in a predictable manner that allows drivers to identify the source of the alarm. UA contexts are transitioning host conflict scenarios where the system issues an alert but the situation resolves itself before the driver needs to intervene. TA contexts represent in-host conflict scenarios where the situation requires the driver to intervene to avoid a collision.

The results suggest that alarm context does matter. Compared to response data that differentiates FA and NA from UA and TA, subjective data shows greater sensitivity and differentiates between all four alarm contexts (FA<NA<UA<TA). Overall, drivers modulate their response according to the driving context not to the presence of an alarm. While drivers evaluated and responded similarly during the FA and NA context, important differences between the groups emerged for the UA and TA contexts.

Younger drivers indicated a high degree of confidence in their own ability across the different conditions. While they adopted a similar response pattern as middle-aged drivers during the TA contexts, these drivers responded less frequently than middle-aged and older drivers during the UA context. Diminished hazard perception ability and the tendency to consider these situations less hazardous likely account for the fewer responses made during these situations by younger drivers.

Older drivers with and without UFOV impairments indicated similar hazard ratings for UA and TA contexts, yet drivers with UFOV impairments responded less frequently in both alarm contexts. Diminished hazard perception ability, slower simple response times, and degraded contrast sensitivity likely account for the fewer and slower responses. Interestingly older drivers with impairments did respond more frequently when warned during the TA context. They also rated FAs and NAs more positively than the other driver groups.

The results of this study suggest applying signal detection theory without concern for the alarm context and driver characteristics is insufficient for understanding how different alarms influence operators and that subjective data can inform design. Researchers are encouraged to combine multiple perspectives that incorporate of both an engineering and human perspective.

Abstract Approved: ________________________________
Thesis Supervisor

____
Title and Department

____
Date

CONTEXT CONDITIONS DRIVERS' DISPOSITION TOWARDS ALARMS

by

Monica Lees

A thesis submitted in partial fulfilment
of the requirements for the Doctor of
Philosophy degree in Industrial Engineering
in the Graduate College of
The University of Iowa

December 2010

Thesis Supervisor: Professor John D. Lee

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH. D. THESIS

This is to certify that the Ph. D. thesis of

Monica Lees

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Industrial Engineering at the December 2010 graduation.

Thesis Committee:
John D. Lee, Thesis Supervisor

Linda Boyle

Jeff Caird

Peter O'Grady

Matthew Rizzo

Geb Thoma

To my family.

ACKNOWLEDGMENTS

I would like to acknowledge my friends, family, colleagues, and mentors for contributing to my journey. I would like to thank John Lee and Matthew Rizzo for challenging me to explore different paths, and for providing insight, guidance, and support when needed. Special thanks to Linda Boyle, Jeff Caird, Geb Thomas and Peter O'Grady for being on my committee and providing valuable comments, and advice throughout the process that has collectively improved the ideas presented here. I would like to thank Jim Sayer and Robert Goodsell from UMTRI and Wassim Najm from the VOLPE Center who provided me with the tools and resources to complete this research. I would like to acknowledge past and present lab members for your friendship, and insights. Finally, to my family who supported me throughout this journey.

ABSTRACT

Collision warning systems represent a promising means to reduce rear-end crash involvement. However, these systems experience failures in the real-world that may promote driver distrust and diminish drivers' willingness to comply with warnings. Recent research suggests that not all false alarms (FAs) are detrimental to drivers. However, very few studies have examined how different alarms influence different driving populations.

The purpose of this research was to examine how younger, middle-aged, and older drivers (with and without UFOV impairments) evaluated and responded to four different alarm contexts – false alarm (FA), nuisance alarm (NA), unnecessary alarm (UA) and true alarm (TA) – when they did and did not receive warnings. FA contexts represent out-of-path conflict scenarios where it is difficult for the driver to identify the source of the alarm. NA contexts represent out-of-path conflict scenarios that occur in a predictable manner that allows drivers to identify the source of the alarm. UA contexts are transitioning host conflict scenarios where the system issues an alert but the situation resolves itself before the driver needs to intervene. TA contexts represent in-host conflict scenarios where the situation requires the driver to intervene to avoid a collision.

The results suggest that alarm context does matter. Compared to response data that differentiates FA and NA from UA and TA, subjective data shows greater sensitivity and differentiates between all four alarm contexts (FA<NA<UA<TA). Overall, drivers modulate their response according to the driving context not to the presence of an alarm. While drivers evaluated and responded similarly during the FA and NA context, important differences between the groups emerged for the UA and TA contexts.

Younger drivers indicated a high degree of confidence in their own ability across the different conditions. While they adopted a similar response pattern as middle-aged drivers during the TA contexts, these drivers responded less frequently than middle-aged and older drivers during the UA context. Diminished hazard perception ability and the tendency to consider these situations less hazardous likely account for the fewer responses made during these situations by younger drivers.

Older drivers with and without UFOV impairments indicated similar hazard ratings for UA and TA contexts, yet drivers with UFOV impairments responded less frequently in both alarm contexts. Diminished hazard perception ability, slower simple response times, and degraded contrast sensitivity likely account for the fewer and slower responses. Interestingly older drivers with impairments did respond more frequently when warned during the TA context. They also rated FAs and NAs more positively than the other driver groups.

The results of this study suggest applying signal detection theory without concern for the alarm context and driver characteristics is insufficient for understanding how different alarms influence operators and that subjective data can inform design. Researchers are encouraged to combine multiple perspectives that incorporate of both an engineering and human perspective.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF DEFINITIONS

Compliance	How the operator behaves when the system indicates a hazard is present (Meyer, 2004). Indicative of whether the driver is willing to respond to alarms.
Dispositional trust	The level of trust upon initial interaction with automation (Merritt & Ilgen, 2008).
Hazard perception ability	The ability to recognize and anticipate hazardous roadway situations. Two main measures are typically collected to evaluate hazard perception ability: response frequency and response time.
Response frequency	The number of times that the driver responded to the situation divided by the number of times the situation occurred.
Response time	The time a warning did or would have occurred subtracted from the time of the driver made a response.
System reliability	Three predominant definitions have been identified within the literature (Sullivan, Tsimhoni, & Bogard, 2008). This dissertation adopts the user perspective in which reliability relates to how the driver subjectively experiences warnings and excludes missed events and adopts the classification of Lees and Lee (2007).
False alarm (FA)	Defined by Lees and Lee (2007) as an alarm associated with a context where the operator is unable to identify the source (e.g. system malfunction).
Nuisance alarm (NA)	Defined by Lees and Lee (2007) as an alarm associated with a context where the operator can identify the source but derives no value.
Unnecessary alarm (UA)	Defined by Lees and Lee (2007) as an alarm associated with a potentially hazardous situation but where the situation resolves itself such that a failure to respond is not associated with a poor outcome (e.g. collision).
True alarm (TA)	Defined by Lees and Lee (2007) as an alarm associated with a hazardous context where the operator must intervene to avoid a poor outcome (e.g. collision).

Self-confidence	The extent to which drivers consider themselves able to handle a particular driving situation.
System appropriateness	The drivers' attitudes regarding whether they consider the alarm provided by the system to be appropriate in the current context.
System trust	Defined by Lee and See (2004) as the drivers' attitudes regarding whether the warning system will help them identify hazards in situations characterized by uncertainty and vulnerability.
System understandability	The drivers' attitudes regarding whether they can understand the source of an alarm provided by the system in a particular context.
System usefulness	The drivers' attitudes regarding whether they consider alarms provided by the system to be of some utility in the current context.
Signal detection theory	A commonly used classification framework for describing performance (Green & Swets, 1966). Four main categories exist within the framework: correct rejection, false alarm, hit and miss.
Correct rejection (CR)	Defined by signal detection theory as a hazard is absent and undetected by the system.
False alarm (FA)	Defined by signal detection theory as a hazard is absent but detected by the system.
Hit	Defined by signal detection theory as a hazard exists and is correctly detected by the system.
Miss	Defined by signal detection theory as a hazard is present but not detected by the system.

CHAPTER 1.

INTRODUCTION

In 2008 alone there were over six million motor vehicle crashes in the United States that resulted in 2.35 million injuries and 37, 261 deaths (NHTSA, 2009). Rear-end crashes constitute approximately one third of all crashes. These crashes result in fewer deaths than other crash configurations but can be costly in terms of injuries and damage. Researchers and manufacturers have begun to explore, develop, and evaluate vehicle technology to reduce rear-end crash involvement.

Collision warning systems have emerged as one promising technological solution to mitigate crash involvement (Dingus, et al., 1997; C. Ho, Reed, & Spence, 2007; Kramer, Cassavaugh, Horrey, Becic, & Mayhugh, 2007; Lee, McGehee, Brown, & Reyes, 2002; Scott & Gray, 2008). However, current systems are imperfect and experience frequent failures (i.e., false alarms (FAs), missed events) that might reduce their effectiveness. Specifically, low collision base rates cause even highly sensitive systems with a high hit rate and a low FA rate to produce a high proportion of FAs (Getty, Swets, Prickett, & Gonthier, 1995; Parasuraman, Hancock, & Olofinboba, 1997; Zabyshny & Ragland, 2003)

A large body of research has demonstrated that high FA rates promote distrust and consequently may diminish an operators willingness to respond to alarms (Bliss, Gilson, & Deaton, 1995; Breznitz, 1983; Wickens & Dixon, 2007). Despite these findings, recent research suggests that not all failures are detrimental to driver trust and compliance (Lees & Lee, 2007). Isolating different alarm types and their influence on trust and compliance may help designers eliminate detrimental alarms while preserving those that provide some utility to the driver. A deeper understanding of how and why different types of alarms lead to driver dissatisfaction is needed.

Unlike other domains, driving involves a heterogeneous population with differing cognitive capabilities, skills, experiences, and personalities. Individual differences may influence how drivers perceive hazards, and how drivers assess their ability to deal with hazards. While important differences may exist between different driving populations that influence the perception of and response to FAs, the systems designed to mitigate crashes assume all drivers are similar. This dissertation examines how driver characteristics influence the perception of and response to different alarm contexts. It does so with the following aims:

- **Aim 1: Examine differences in self-confidence and hazard perception between different driving populations.** The purpose is to examine how drivers assess hazards when they do and do not receive collision warning system alarms. Specifically, this aim determines how younger, middle-aged, and older drivers 1) assess their ability to handle hazardous situations and 2) evaluate hazards within different contexts. Older drivers are further differentiated into normal older drivers and drivers with cognitive impairment characterized by useful field of view (UFOV) deficits.
- **Aim 2: Assess how different driving populations perceive different alarm types**. The results will be used to understand how drivers perceive alarms that occur within different driving contexts and how different alarms influence driver trust.

Chapter 2 provides an overview of the literature relevant to the overall scope of this dissertation. An overview of the primary causes of rear-end collisions is provided as well as the potential benefits of collision warning system technology. The remainder of the background focuses on the difficulty in measuring reliability and different perspectives on characterizing FAs. Individual driver differences are also examined in terms of hazard assessment, self-confidence, and driving style. These factors are examined in terms of the potential benefit derived from using collision warning systems, and the consequences different alarm contexts might have on different driver groups. Chapter 3 provides the methodology for the experiment used to address the aims of this dissertation. Chapter 4 provides the results and conclusions relevant to aim 1 by examining how different driver groups respond to different alarm contexts. Chapter 5 provides the results and conclusions relevant to aim 2 by examining how different alarms influence trust and driver perceptions. Chapter 6 discusses important findings relevant to the literature, provides a summary of the key findings and limitations of the study.

CHAPTER 2.

DRIVER RESPONSE TO IMPERFECT COLLISION WARNING SYSTEM TECHNOLOGY

Rear-end crashes account for approximately one third of all crashes (NHTSA, 2006). While they do not result in a large number of fatalities, rear-end crashes represent a significant concern in terms of immediate and ongoing costs (e.g., congestion, loss of productivity, medical and rehabilitation costs). As such, there have been efforts to 1) understand what factors contribute to these crashes and 2) to design and evaluate different methods, such as collision warning systems, for mitigating rear-end crash involvement. However, the design of such systems is complicated by the low base rate of collisions for the individual. These systems may produce a high number of false alarms (FAs) that can result in driver distrust and result in drivers delaying or inhibiting responses when warned. At the same time research suggests that not all alarms are detrimental to users. Certain individual and contextual factors are likely to influence how such systems benefit drivers. The overall goal of this research is to better understand how younger, middle-aged and older driver perceive and respond to different types of collision warning system alarms.

<u>What behaviours lead to rear-end crashes?</u>

Rear-end crashes have been attributed to drivers adopting unsafe following distances and to driver inattention (Dingus, et al., 1997; Knipling, et al., 1993). Two factors may contribute to the tendency of drivers to engage in unsafe following behaviour. First, drivers develop expectations about typical car following conditions. Over time drivers learn that large changes in relative velocity rarely occur when following another vehicle (Dingus, et al., 1997). Second, drivers are often able to detect changes in headway distance and compensate accordingly (Dingus, et al., 1997). Drivers can accurately judge the direction of relative velocity—whether the gap between their own vehicle and another is increasing or decreasing (Mortimer 1988), by using visual cues associated with the angle subtended by the lead vehicle (LV) and the rate of change of this angle to modulate their braking to avoid a collision (Mortimer, 1990). However, drivers perceive these cues imperfectly and perceptual thresholds can delay the detection of a LV that is slowing (Mortimer, 1988; Park, Lee, and Koh, 2001). Drivers also underestimate the rate of LV deceleration, especially when deceleration is high (Park, Lee, and Koh, 2001). Such factors may result in drivers failing to cope when confronted with a vehicle at rest or a vehicle with a high rate of deceleration (Dingus, et al., 1997).

Although these perceptual failures are important, inattention appears to play a more predominant role in these crashes. For example, in one study inattention contributed to over 60% of rear-end collisions (Knipling, et al., 1993). Epidemiological studies provide additional evidence that using in-vehicle devices such as cell-phones while driving substantially increases crash risk (Redelmeier & Tibshirani, 1997). A meta-analytical study used data from 33 studies to evaluate the effects of cell phones on driver performance (Caird, Willness, Steel, & Scialfa, 2008). The study provided substantial evidence that using a cell-phone while driving can slow driver responses to objects and events. Event/stimulus, driver age, and task (e.g. conversation, cognitive task) appear to moderate such delays. For example, the degradation in response ranged from 0.17 s for simple detection tasks to 0.36 s for LV braking events. Comparisons between distracted and non-distracted driving conditions for younger and older drivers suggest that distraction may be more detrimental to older drivers. Specifically, distraction tasks delayed younger driver responses by 0.19 s but delayed older driver responses by 0.46 s. Other devices likely induce similar impairments, reducing the driver's capacity to anticipate and adequately compensate when hazards arise. For example, one study found that a speech based email system delayed driver responses to a periodically braking LV by 0.31 s (Lee, Caven, Haake, & Brown, 2001). These and other studies demonstrate that distraction-related delays and inattention to the LV are powerful contributors to rear-end crashes.

Several countermeasures have been introduced to reduce crash involvement, yet some of these failed to benefit drivers because they did not adequately account for factors shown to increase crash likelihood. For example, advance brake warnings activate the LV brake lights when the driver rapidly releases the accelerator and are meant to provide the following vehicles driver more time to respond to rapid rates of deceleration. However, to be effective the driver must be looking at the vehicle ahead. A Monte Carlo simulation demonstrated the benefits of such systems especially when the driver was alert and the intervehicle headway was less than 1.00 s (Shinar, Rotenberg, & Cohen, 1997). However, in a field operational test the technology failed to reduce rear-end crashes suggesting that the conditions for such systems to be effective (e.g., driver attends to the LV, driver does not delay response) may not be met in the real world (Shinar, 2000). Such results suggest that drivers might benefit most from warnings that direct attention to the road rather than warnings that alert drivers of situations with a high rate of closure.

Mitigating rear-end crash involvement with collision warning systems

Collision warning systems represent a promising means to reduce rear-end crash involvement—not by compensating for perceptual limits but by redirecting the drivers' attention. Such systems direct the drivers' attention to impending hazards through an alarm (e.g., flashing lights, feedback from the accelerator, seat vibrations) when some threshold based on velocity and distance is exceeded. In 1992 Daimler-Benz researchers estimated that 60% of rear-end crashes could be prevented if drivers had 0.50 seconds more to respond; increasing this time to one second, might prevent 90% of all rear-end crashes (Ankrum, 1992). Estimates obtained through Monte Carlo simulation suggest that such systems would reduce crashes from between 37 to 74% (Farber & Paley, 1993; Knipling, et al., 1993; Najm, Wiacek, & Burgett, 1998).

Although the estimated benefit of such systems is somewhat variable, a number of simulator and on-road studies suggest that such systems offer great potential in helping drivers manage and avoid potential threats while driving (Dingus, et al., 1997; C. Ho, et al., 2007; Kramer, et al., 2007; Lee, et al., 2002; Scott & Gray, 2008). These studies suggest that perfectly reliable collision warning systems can benefit drivers by increasing headway time (Dingus, et al., 1997), reducing response times to hazardous situations (C. Ho, et al., 2007; Kramer, et al., 2007; Lee, et al., 2002; Scott & Gray, 2008) and reducing collision involvement (Kramer, et al., 2007; Lee, et al., 2002). These systems have been shown to be beneficial even when drivers are distracted (Lee, et al., 2002). For example, one study found that alarms that redirect attention diminished response time delays associated with distraction by 0.11 s to 0.86 s depending on the thresholds used to trigger the warning (Lee, et al., 2002). Although very few of these studies have examined older drivers, one study did find that older and younger drivers derived similar benefits when interacting with perfectly reliable collision warning systems (Kramer, et al., 2007). Despite these promising findings, such systems experience failures in real world settings that may jeopardize such benefits.

Early warnings are most valuable to the driver but compromise system reliability

While an ideal collision warning system would be 100% reliable, noisy data and great uncertainty within the driving environment make this unlikely (Getty, et al., 1995; Parasuraman, et al., 1997). According to Allstate Insurance (2008), the average driver in the United States will experience a crash every 10 years. This means that while a large number of crashes occur nationally, collisions for the individual are an infrequent event that must be correctly identified by the system (Parasuraman, et al., 1997).

At the same time, the largest safety benefits occur when drivers receive early warnings (Abe & Richardson, 2004, 2005, 2006a, 2006b; Brown, Lee, & McGehee, 2001; Lee, et al., 2002; Scott & Gray, 2008). This creates a paradox in that the conditions needed to aid drivers can also cause the system to produce errors. More FAs (the system indicates a hazard is present when there is not) occur because these systems utilize an extremely low threshold to provide sufficient time for the driver to respond and to minimize the more costly error of missing true events that have an extremely low frequency (Parasuraman, et al., 1997).

The concept of positive predictive value (PPV), the probability that an alarm will indicate a true hazard, further demonstrates this (equation 3, Bustamante, Bliss, & Anderson, 2007; Getty, et al., 1995; Parasuraman, et al., 1997; Zabyshny & Ragland, 2003). PPV takes into account the base rate of having a rear end collision, P(Hit). In the current example (previously used by Zabyshny and Ragland, 2003) it is assumed that P(Hit) is 0.002 and P(Correct Rejection) = 0.998. In such a case even if a collision warning system has a high hit proportion (HP, equation 1) of 0.99 and a low FA proportion (FAP, equation 2) of 0.001 the PPV will be 0.17. That is, even when the system is extremely sensitive the low base probability of having a collision causes the driver to receive approximately five FAs for every TA.

$$\textit{Hit Proportion (HP)} = \frac{\text{Hits}}{\text{Hits} + \text{Misses}} \tag{1}$$

$$\textit{False Alarm Proportion (FAP)} = \frac{\text{False Alarms}}{\text{False Alarms} + \text{Correct Rejections}} \tag{2}$$

$$\textit{Positive Predictive Value (PPV)} = \frac{\text{HP} \times \text{P(Hit)}}{[\text{HP} \times \text{P(Hit)}] + [\text{FAP} \times \text{P(Correct Rejection)}]} \tag{3}$$

Figure 1 represents the collection of evidence over time by a system and by a driver when determining whether a response is required to avoid a collision within a particular driving context. Two situations are illustrated: a) in the top figure the situation requires the driver to brake, and b) in the bottom figure the situation will resolve itself regardless of whether the driver responds. In both situations the system utilizes a more conservative threshold (requires less evidence) and consequently will determine that action is necessary (C_S) prior to the driver (C_D). This difference in thresholds influences system performance and how the driver might perceive warnings.

Earlier warnings allow the driver with more time to monitor the situation and determine an appropriate response. At the same time these alarms have a lower predictive power (will produce more FAs). The greater the separation between C_S and C_D, the greater discrepancy there might be regarding the presence of a hazard. For example, if warnings are provided too early the hazard might resolve itself (Figure 1b) or as demonstrated by Figure 1a if warnings are provided too early the driver may be unable to match the source of the alarm to the current driving context (Abe & Richardson, 2004). Driver trust may diminish in situations where the driver cannot corroborate the system's assessment, especially if the driver cannot develop expectations about how the system is functioning.

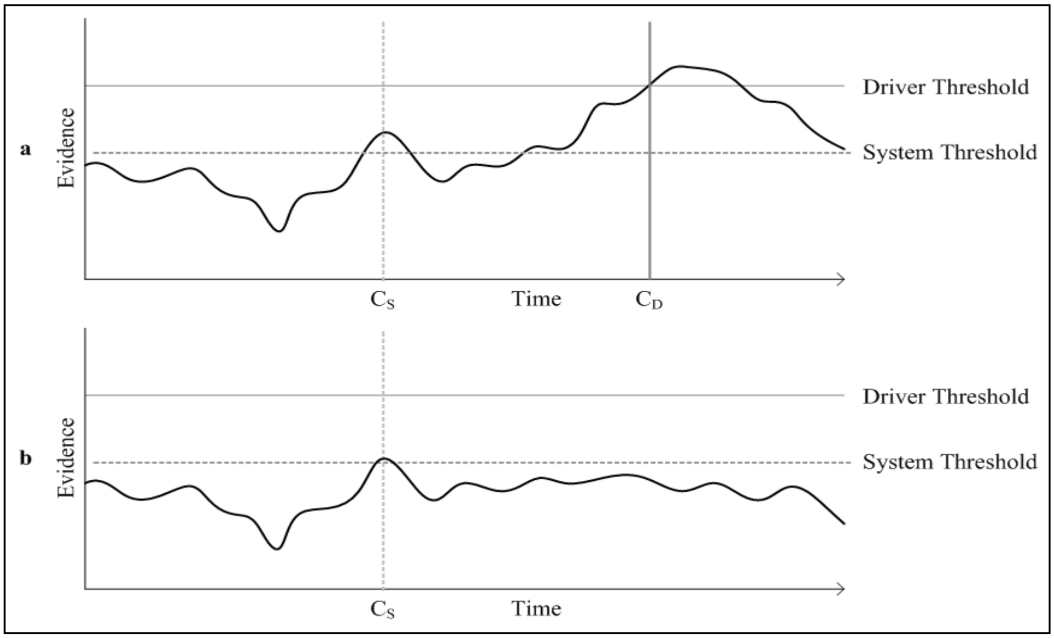

Figure 1. An example of how the system and driver may utilize different thresholds for evidence when determining threats. C_S and C_D represent the systems or drivers decision that a hazard exists and action is necessary to avoid a collision. In figure 1a, the situation will require the driver to respond to avoid a collision, whereas the hazard resolves itself without the driver intervening in figure 1b.

Simple measures of reliability confound failure type and frequency

Collision warning system reliability is somewhat difficult to define and several perspectives have been adopted (Sullivan, et al., 2008). Sullivan, Tsimhoni and Bogard (2008) highlighted three dominant perspectives. From an engineering perspective, reliability relates to the ability of the system to produce consistent results under similar conditions. From a functional perspective, reliability relates to how many errors the system generates (missed events and FAs) per unit of time. Finally, from a user perspective reliability relates to how the driver subjectively experiences warnings and excludes missed events. Different perspectives have led to a diverse set of reliability measures, including the number of errors over time, the FA rate or some combination of cue-probability and FA rate.

Research examining the effects of unreliable automation has primarily used a functional perspective. For example, Getty et al. (1995) examined operator responses across five levels of PPV (the proportion of TAs was 0.25, 0.39, 0.50, 0.61 and 0.75). While the analysis was only based on three subjects, the study found evidence that response times were reduced when the PPV was greater than 0.50.

Wickens and Dixon (2007) used data from 20 studies to identify the value at which automation (relative to performance when no automation is present) fails to benefit the operator or becomes a drawback because of its level of reliability and whether such a value is influenced by task workload. The study found that when reliability falls below 0.70, performance (e.g., reaction time, accuracy) degrades so much that the operator would be better performing the task without the automation and that operators depend on automation more for high workload tasks. These findings only provide a rough estimate of when automation might become a hindrance to the operator. Several factors such as expectations about system reliability, and understanding the source of the failure might influence the value for this threshold (Lee & See, 2004) and were not considered in the analysis.

As mentioned before, this functional approach fails to distinguish failure types (Ben-Yaacov, Maltz, & Shinar, 2002; Donmez, Boyle, & Lee, 2006; Maltz & Shinar, 2004). Reliability combines HP and FAP so that systems with very different performance profiles have the same reliability. This perspective of reliability fails to describe the frequency with which a driver experiences different failures. This point is demonstrated graphically in the Receiver Operating Characteristic (ROC) curves plotted in Figure 2. The ROC plots the probability of a hit, P(Y/SN), versus the probability of a FA, P(Y/N) for all response criterion (C) values when sensitivity (d') is held constant (a more detailed description of signal detection theory is provided in the next section). For the line originating at (0.2, 0), the overall reliability of the system is 20%. The different points represent unique combinations of failures that a driver can experience under this level of system reliability. In one extreme (A), all failures occur when the system fails to inform the driver of a hazard (miss). In the other extreme (C), all failures occur when the system incorrectly informs the driver that a hazard exists (FA). At B, the system experiences an equal number of the two failure types. For points prior to B, the number of missed events exceeds the number of FAs. In points after B, the number of FAs exceeds the number of missed events. Thus even when the system has the same level of reliability drivers may not experience the same proportion of misses and FAs. This is important given that these different failures can have different consequences for a person operating in tandem with automation (Abe & Richardson, 2006a; Dixon, Wickens, & McCarley, 2007). As a result, adopting a functional perspective of reliability makes it difficult to determine if the behaviour observed is due to the overall reliability of the system, or a specific type of failure.

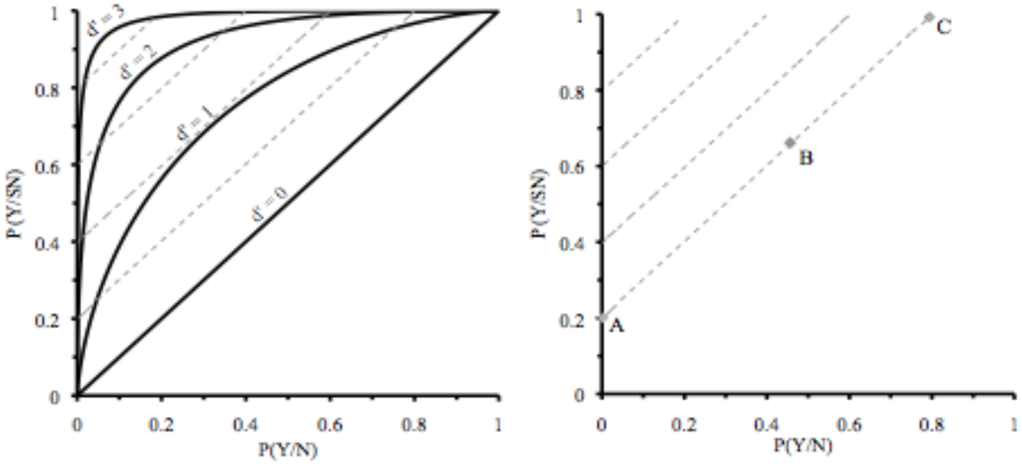

Figure 2. Receiver Operating Characteristic (ROC) curves for d' values of 0, 1, 2, and 3. The ROC plots the probability of a hit, P(Y/SN), versus the probability of a FA, P(Y/N), for all values of C for a given value of d'. The right plot provides overall reliability, 1- P(Y/SN) + P(Y/N), equal to 0.2, 0.4, 0.6, and 0.8. For reliability of 0.2, point A represents instances where all failures are missed events, 1- P(Y/SN), point B represents instances where the probability of a missed event is equal to the probability of a FA, and point C represents instances where all failures are FAs, P(Y/N).

<u>Signal detection theory: FAs versus missed events</u>

Signal detection theory represents a commonly used classification framework for describing system and human performance. Figure 3 graphically illustrates the main concepts of signal detection theory as presented by Sorkin and Woods (1985). Discrimination of noise (N) from signal-plus-noise (SN) “depends on the probability density distribution of the statistic Z” (Sorkin & Woods, 1985, p. 53). f(Z/N) represents the probability of obtaining a given Z value when only N is present; f(Z/SN) represents the probability of obtaining a given Z value when SN is present (Sorkin & Woods, 1985). Two parameters influence discrimination: 1) sensitivity (d’) and 2) response criterion (C). d’ represents the separation between the two distributions. C represents the amount of evidence required by the decision maker to report that a signal is present.

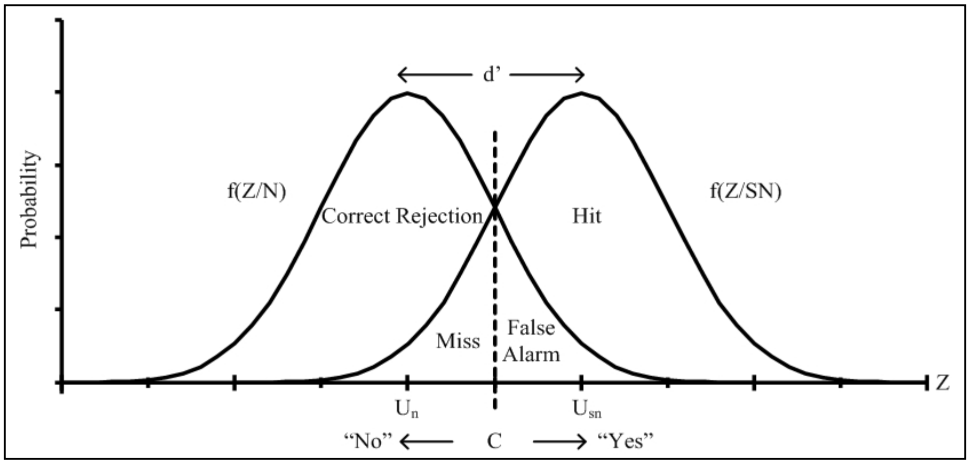

Figure 3. A graphical illustration of signal detection theory based on Sorkin and Woods (1985).

According to signal detection theory four outcomes are possible: 1) *Hit*: a hazard exists and is correctly detected, 2) *Correct Rejection*: a hazard is absent and not detected, 3) *FA:* a hazard is absent but incorrectly detected, and 4) *Miss*: a hazard is present but not detected. Figure 4 shows how these outcomes are influenced by changes in d'. Increasing d' decreases the degree of overlap between the two distributions such that the hit rate increases and the FA rate decreases. Although it is ideal to maximize separation between N and SN distributions, d' is typically limited by the hardware and algorithms used by the collision warning system (Sorkin & Woods, 1985).

Figure 5 shows how discrimination is influenced by changes in C. A hazard is considered to be present when $Z \geq C$. When C moves towards the SN distribution the hit rate and FA rate decrease because of an increased propensity to respond that a signal is absent. In contrast, when C moves toward the N distribution the hit and FA rate increase because of an increased propensity to respond that a signal is present. The criterion used by collision warning systems is often adjusted to reduce missed events at the cost of an elevated FA rate.

While signal detection theory offers a useful tool for evaluating tradeoffs in system performance there are important limitations to this representation relevant to decision making in complex environments. For example, while signal detection theory assumes decision making occurs in a static environment (Balakrishnan, MacDonald, Busemeyer, & Lin, 2007) this may not be true in driving where situations evolve over time. Figure 6 provides a graphical example of how both d' and the variation in the N and SN distributions might change over time as the decision maker accumulates additional evidence (Balakrishnan, et al., 2007).

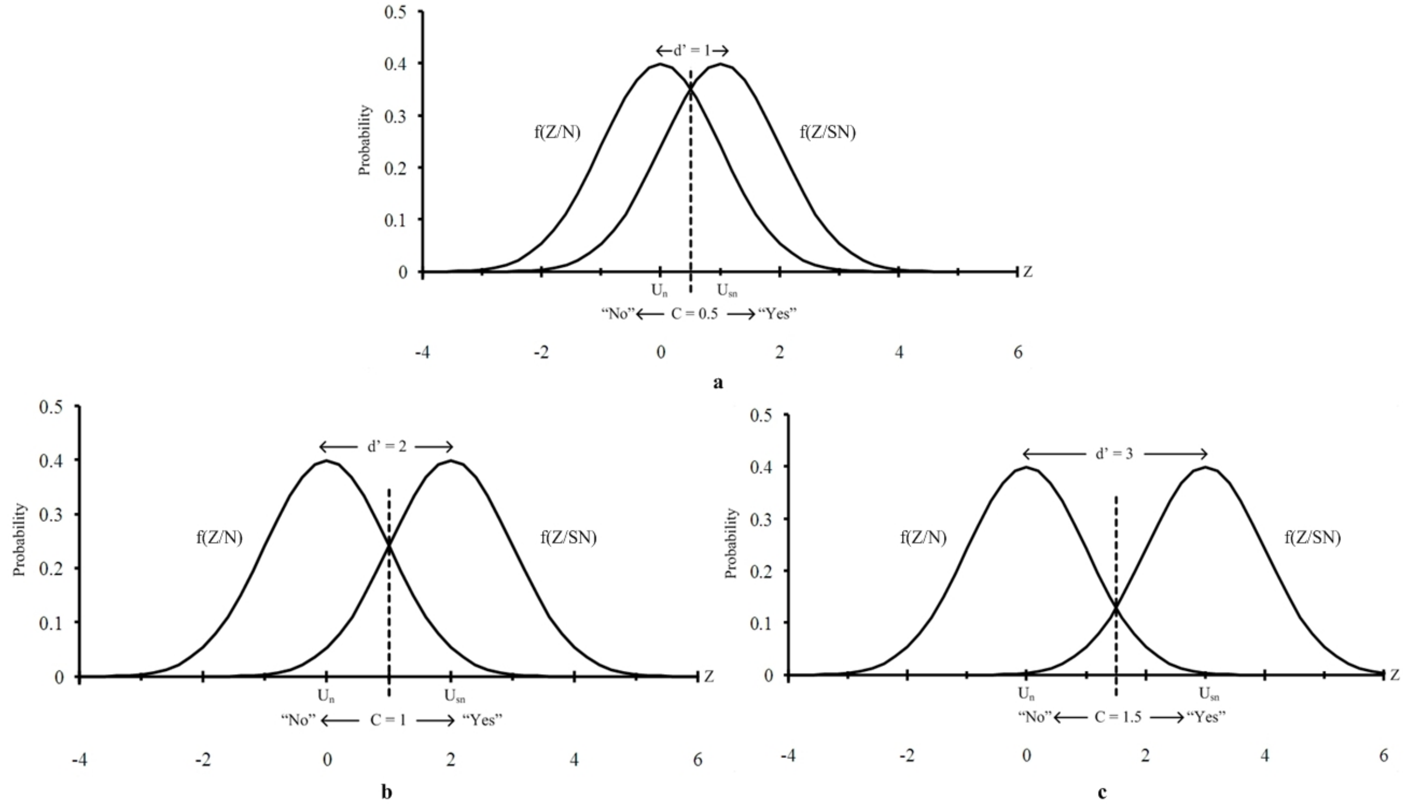

Figure 4. A graphical illustration of how changes in sensitivity (d') influence the hit and FA rate (Sorkin & Woods, 1985). The figure illustrates the probability distributions of Z when a) d'=1, b) d'=2, and c) d'=3. As d' decreases, the hit rate will increase and the FA rate will decrease.

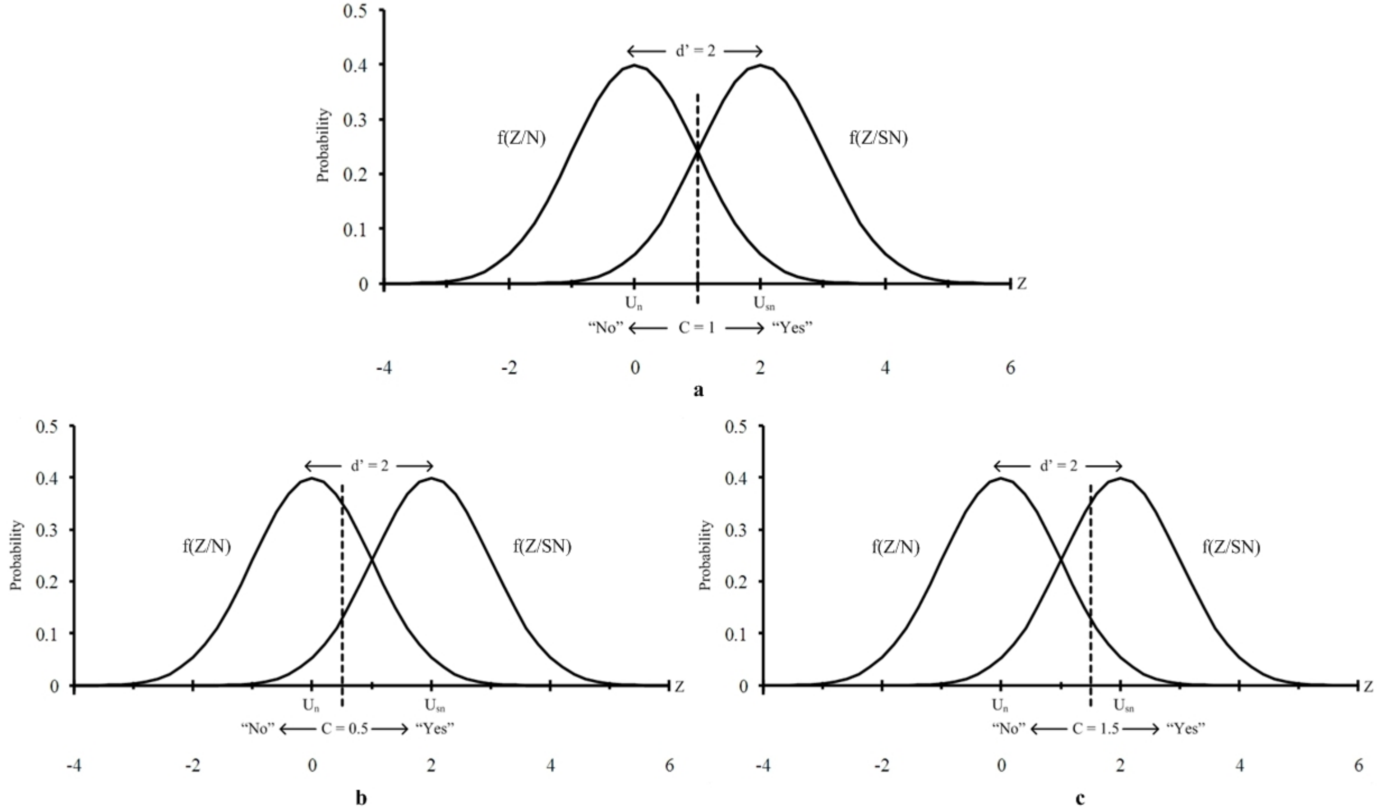

Figure 5. A graphical illustration of how changes in C influence the hit and FA rate for cases in which d'=2 (Sorkin & Woods, 1985). As C moves towards the noise (N) distribution both the hit rate and FA rate increase (figure 5b). As C moves towards the signal-plus-noise (SN) distribution both hit rate and FA rate decrease (figure 5c).

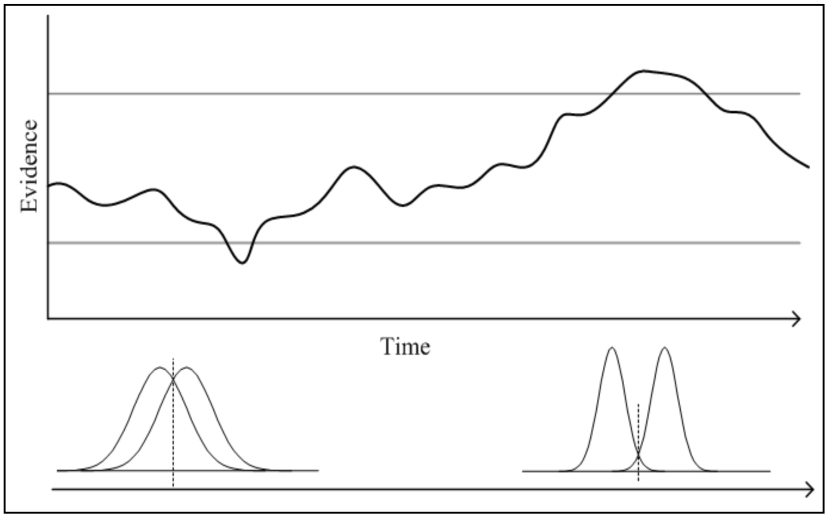

Figure 6. A graphical illustration of how the N and SN distributions might change over the course of an event.

By failing to consider decision making over time, the signal detection theory framework fails to account for how the driver might influence performance of the system. As mentioned previously, with early warnings the driver has the capacity to change the environment such that the threat no longer materializes. The warning may be false because the system does not have all the evidence and because the actions of the driver can change the situation. In such cases, delaying the warning might eliminate FAs. Balakrishnan et al. (2007) developed a dynamic signal detection theory to account for these limits.

In summary, signal detection theory represents a widely used framework for describing and evaluating system performance. The theory identifies two broad failure categories, FAs and missed events that might influence how drivers interact with collision warning systems. However, signal detection theory may be too simplistic for evaluating collision warning system performance. Other researchers have extended this framework to better evaluate such systems. For example, some extensions have been conducted to make classical signal detection theory more conducive to considerations of time (Allendoerfer, Pai, & Friedman-Berg, 2008; Balakrishnan, et al., 2007), and to joint decision making (Sorkin & Woods, 1985). Although these extensions make signal detection theory more useful in describing how system failures influence operator responses they do not address the reason why FAs undermine trust and compliance with the system.

Why might FAs diminish the benefits of a system

FAs are likely to be a prominent issue for collision warning systems because the design of these systems emphasizes the need to eliminate or minimize missed events. Data from a recent field operational test provides evidence that collision warning systems can generate a large number of FAs (NHTSA, 2005a). Over the course of the study, a number of algorithms were initially tested for implementation in a small fleet of vehicles. The algorithm was subjected to three major refinements to reduce FAs. Despite these refinements, drivers received 0.62 alarms per 100 km of travel with the final algorithm (algorithm C), and according to one evaluation 97% of these were false (Najm, Stearns, Howarth, Koopmann, & Hitz, 2006).

Generally speaking, FAs have been shown to influence trust and compliance—how the operator behaves when the system indicates a hazard is present (Dixon & Wickens, 2006; Donmez, et al., 2006; Meyer, 2004). FAs may result in a “cry wolf ” effect where the likelihood of the driver responding to a valid warning is reduced, or the driver delays response for long periods as the source of a particular alarm is investigated (Bliss & Acton, 2003; Bliss, Dunn, & Fuller, 1995; Breznitz, 1983; Hagenouw, 2007;

Kestin, Miller, & Lockhart, 1988; Sorkin, 1988; Xiao & Seagull, 1999). However, trust may recover under certain conditions. For example, some studies have found trust can recover if the system only experiences a small number of errors (Gao & Lee, 2006; Lee & Moray, 1992, 1994).

Some evidence suggests that people match their frequency of response to the expected probability of true warnings (Bliss & Acton, 2003; Bliss, Gilson, et al., 1995; Getty, et al., 1995). For example, Bliss and Acton (2003) found that an increase in FAs (25% or 50% of alarms were false) resulted in drivers responding less frequently and less correctly when attempting to avoid a collision as compared to drivers who only received true alarms (TAs; 0% of alarms were false). Even in cases where the operator continues to respond to warnings, responses may be slower and less accurate (Dingus, et al., 1997; Getty, et al., 1995; Grounds & Ensing, 2000). Either outcome can diminish the benefits of the system.

Several factors might influence how FAs influence the operator. For example, evidence suggests that trust and performance may recover from errors when the system provides accurate information or when operators learn to accommodate for failures (Abe & Richardson, 2006a; Breznitz, 1983; Kantowitz, Hanowski, & Kantowitz, 1997; Lee & Moray, 1992). Kantowitz et al. (1997) found that trust diminished on inaccurate trials but subsequently increased when the system provided accurate information. Trust diminished to a greater extent when multiple failures occurred in subsequent trials before the system provided accurate information.

Abe and Richardson (2006) investigated how alarm timing influenced driver responses and attitudes toward collision warning system failures (missed events and

FAs). Drivers were exposed to 11 potential collision situations. Trials were set up in the following sequence: 1) two no alarm trials, 2) three TA trials, 3) one FA trial, 4) two TA trials, 5) one missed event trial, and 6) two no alarm trials. Drivers were assigned to either an early or late warning timing. Trust and response times were examined over the different types of trials (TA trials, trials before a FA, FA trial, TA trials after FA, and missed event). Drivers presented with early alarms responded faster and trusted the system more than drivers presented with late alarms. Drivers presented with late alarms did not delay responding when the system missed an event. For drivers presented with late alarms, trust did not differ across trails. Drivers presented with early alarms were more likely to respond when receiving a FA and to delay a response when the system missed a hazard. For these drivers trust diminished on the FA trial but subsequently recovered on TA trials after the FA.

In addition, recent research suggests that not all FAs are detrimental to trust and compliance (Lees & Lee, 2007). Lees and Lee (2007) examined two types of alarm failures (FAs and unnecessary alarms - UAs). FAs represent random collision warning system activation. The random nature of such alarms makes it difficult for drivers to understand or predict their occurrence. In contrast, UAs represent activation of the collision warning system in situations judged hazardous by the algorithm. However, drivers may consider these alarms to be unnecessary. The context associated with these alarms can potentially help the driver understand both the source of the failure and how the system operates. The study found that unlike FAs, the context associated with UAs fostered system trust and compliance that carried over into situations that required driver intervention. The authors argued that the context associated with UAs may have caused drivers to incorporate this information when assessing future threats, making them more cautious and sensitive to potential hazardous situations. These findings suggest that the context and cause of FAs matter and may influence the how drivers perceive the system.

Differentiating types of FAs from a user's perspective

In the context of collision warning systems, the label "FA" embodies a variety of failure contexts. For example, while the study by Dingus et al. (1997) paired FAs with 1) road signs, 2) gradients, 3) the driver carrying out a turn, other researchers have examined FAs devoid of any context (Ben-Yaacov, et al., 2002; C. Ho, Reed, & Spence, 2006; Lees & Lee, 2007; Maltz & Shinar, 2004). Under signal detection theory the failure context does not matter and consequently these failures are treated equally. Yet receiving a warning when driving on an empty road likely evokes a different response than that of a warning received when a LV brakes to turn.

Figure 7 depicts different alarm contexts that occurred during the Automotive Collision Avoidance Field Operational Test (ACAS FOT) (NHTSA, 2005a). The circle on each image represents the target that triggered the imminent collision warning system alarm. According to signal detection theory all of these warnings represent a FA. However, because these warnings occur in fundamentally different contexts they may not have a uniform effect on the driver. For some of the contexts it may be easy for the driver to understand why the warning has occurred (Figure 7b-d) and in some cases the warning may appear random (Figure 7a). For some cases, the driver may agree a hazard exists (Figure 7c-d) and in some cases the warning may seem unwarranted (Figure 7a-b). It seems possible that information may be lost if all of these failures are lumped together into a broad "FA" category.

Figure 7. **Different contexts that a driver might receive forward collision system warnings**

Along these lines, several researchers have suggested that signal detection theory is inadequate to describe and evaluate how and why failures influence the operator. As a result, several alternative frameworks have been proposed to better understand the influence of alarm context on operators (Allendoerfer, et al., 2008; Barnes, Gruntfest, Hayden, Schultz, & Benight, 2007; Friedman-Berg & Allendoerfer, 2008; Lees & Lee, 2007; Seagull & Sanderson, 2001; Woods, 1995).

Some researchers have adopted an engineering perspective that evaluates how drivers respond to alarms occurring within different conflict scenarios. For example, one evaluation of the field operational test conducted on ACAS in 2005 examined imminent warnings (for Algorithm C) within different conflict scenarios (NHTSA, 2005a, 2005b). Three broad conflict scenarios were identified 1) in-host path (IHP – 29% of alarms), 2) transitioning-host path (THP – 35% of alarms), 3) out-of-host path (OHP – 36% of alarms). IHP scenarios are cases where the system is triggered by a vehicle that remains within the driver's lane (e.g. lead vehicle braking, stopped vehicle, car following). THP scenarios are cases where the system is triggered by a vehicle in front of the driver, but the situation resolves itself with one or both vehicles making a lateral manoeuvre. OHPs are cases where the system is triggered by an object outside of the drivers' lane (e.g. signage, overhead bridge). In 64% of OHP and 67% of THP scenarios the driver did not respond to the alarm or situation (no braking or steering). In contrast, drivers responded 80% of the time during IHP scenarios. Even though braking was more common in IHP scenarios, brake response times were similar across IHP and THP scenarios. At the same time drivers adopted similar responses when they received warnings as compared to when warnings were muted across the different conflict scenarios. Subsequently, the report concluded that 60% of the imminent warnings were valid. This estimate of TAs is much higher than that reported by Najm et al. (2006) which used the same data.

Other researchers have extended signal detection theory to incorporate more categories. Allendoerfer, Pai, and Friedman-Berg (2008) discussed how signal detection theory could be expanded to include the interaction between automation and operator. A variety of factors are proposed to influence the alarm signal, the operators trust in an alarm, and the operators assessment of the signal or decision criterion (Allendoerfer, et al., 2008). For example, errors in radar data, and manoeuvre type might influence input data (evidence available), whereas workload and urgency might influence the criterion used by the operator. When event timing is considered two additional signal detection theory categories emerge: 1) Hit But Late (HBL) and 2) FA But Early (FABE). HBLs occur when the controller has not taken action when a TA activates, and the hazardous situation has been detected too late for the controller to avoid an error. FABEs occur when the controller has not taken action when the alarm activates, but the operator does not trust or agree with the alarm because they have additional information not incorporated by the system. In such instances, the systems long look ahead time prevents the system from incorporating incoming information such as the conflict being resolved. If the system was given additional time to process incoming information, the alarm would not have occurred. Just like tradeoffs between missed and FAs there exist tradeoffs between HBLs and FABEs where longer lead times result in more FABEs but fewer HBLs.

Others have developed more general frameworks for understanding different types of alarms. For example, Xiao and Seagull (1999) identified three major alarm types 1) FAs, 2) nuisance alarms (NAs), and 3) inopportune alarms. FAs were attributed to system unreliability, and faulty calibration or connection of monitoring devices. NAs were defined as lacking value to the human operator because the critical state indicated does not apply within the current context. Finally, inopportune alarms were defined as warnings that occur at the wrong time. These alarms often co-occur with other alarms but represent a minor disturbance. One study found that even when the overall system reliability is high such interruptions degrade operator trust and performance (Parasuraman & Miller, 2004).

Alternatively, Woods (1995) outlined a general framework for possible failure modes 1) nothing is wrong, the context does not matter 2) within the current context nothing is wrong 3) nothing is wrong because the system state was induced by the operator 4) the operator is already aware of the problem, 5) the operator does not need to know about the problem and 6) the problem is inconsequential. Friedman-Berg and Allendoerfer (2008) used a similar categorization of NAs produced by two systems used to aid air traffic controllers. NAs were identified as those in which the controller did not respond to the alarm and no operational error occurred. They identified seven categories for application in the air traffic control domain: 1) no action necessary, 2) already addressing it, 3) someone else's problem, 4) obnoxious, 5) using other types of separation, 6) repeat alarm, 7) surveillance or tracking error (Friedman-Berg & Allendoerfer, 2008). In the medical domain, a recent study identified four categories of alarms based on the type of operator response (Seagull & Sanderson, 2001). The four types of alarms included: 1) the alarm is unexpected and the operator takes corrective action or makes a change, 2) the alarm is expected or intended and consequently not corrected, 3) the alarm is invalid and ignored, 4) the alarm is a reminder to initiate an action.

These alternative classifications further differentiate alarms to better understand their influence on operators. Classifications that use an engineering perspective focus on how drivers respond within constrained conflict scenarios. Other classifications differentiate alarms from a user perspective and infer rationale for why operators respond differently to alarms. Regardless of the focus, such classifications offer a deeper understanding of how and why different types of alarms might contribute to dissatisfaction and may allow designers to develop better algorithms. Expanding the classification can help isolate different alarm types and their influence on compliance and trust. Improved understanding of different types of alarms may allow designers to develop algorithms that eliminate detrimental alarms while preserving those that provide some utility.

Using dimensions of trust to better understand FAs

Noting the limitations of signal detection theory, Lees and Lee (2007) offered an alternative classification based on trust. Trust has been shown to mediate reliance between people, and between people and automation (Abe & Richardson, 2005, 2006a; de Vries, Midden, & Bouwhuis, 2003; Gupta, Bisantz, & Singh, 2002; Lee & Moray, 1992, 1994; Lee & See, 2004; Muir, 1987; Rempel, Holmes, & Zanna, 1985). Trust describes the drivers attitude regarding whether the system will help them identify hazards in situations characterized by uncertainty and vulnerability (Lee & See, 2004).

Three factors have been proposed by Lee and Moray (1992) to influence trust in automation: purpose, performance, and process (Lee & Moray, 1992). *Purpose (intent)* defines the reason the automation exists, and refers to the designers' intent regarding what information the system should provide and when/how the system should warn the driver. *Performance (utility)* defines what the automation does, and refers to the ability of the system to aid the driver in collision avoidance. *Process (predictability)* defines how the collision warning system works, and is governed by the systems algorithms, sensors and alarm logic. These three dimensions suggest that drivers may view the behaviour of a collision warning system differently than might be expected from the description of warnings based on signal detection theory. That is, these dimensions consider both the users' and the designers' perspectives.

Table 1 presents a range of alarm types identified using these dimensions (Lees & Lee, 2007, 2009). This theoretical framework requires further examination to identify

Table 1. A sample of collision warning system alarms based on the performance, process, and purpose dimensions of trust that could infleunce the users' perception of the system (Lees & Lee, 2007, 2009).

Alarm Type	Purpose, Process and Performance		Example	Consequences
True (Accurate)	Intended Predictable Useful		An alarm associated with a hazardous driving context that the driver might not otherwise avoid.	May enhance driving performance and trust.
Unnecessary	Intended Predictable	Non-Useful	An alarm associated with a situation judged hazardous by the designer, but not by the driver. The driver can understand what triggered the alarm.	May help drivers understand how the system works, but may annoy drivers if frequent.
Incomprehensible	Intended Useful	Unpredictable	An alarm associated with a hazardous driving context that the driver might not otherwise avoid, but is not recognized as such by the driver.	May diminish trust and compliance
Unappreciated	Intended	Unpredictable Non-useful	An alarm that associated with a situation judged hazardous by the designer, but is not understood or appreciated by the driver.	May diminish trust and compliance.
Fortuitous	Predictable Useful	Unintended	An alarm that is inconsistent with the stated purpose of the system, but which helps the driver avoid hazards.	May enhance driving performance and trust, but could lead drivers to use the system differently than the designer intended.
Inadvertent/ Nuisance	Predictable	Unintended Non-Useful	An alarm triggered by events that do not pose a threat to the driver and were not intended by the designer, such as vehicles in the adjacent lane, roadside objects, and clutter on curves (Zador, Krawchuk, & Voas, 2000).	May help drivers understand how the system works, but may undermine system credibility.
Unforeseen	Useful	Unintended Unpredictable	An alarm triggered in manner that is inconsistent with the designers intent and is not understandable to the driver, but is useful in avoiding a hazard.	May enhance driving performance and trust, but could lead to drivers to use the system differently than the designer intended.
False		Unintended Unpredictable Non-useful	An alarm triggered by sensor noise, system malfunction that neither helps the driver understand the system or respond to threats.	May diminish trust and compliance.

which dimensions are most central in promoting trust and compliance with collision warning systems. Four of the alarm types highlighted in table 1 (UAs, NAs, FAs, and TAs) are examined in this dissertation.

Individual characteristics may influence the perception of alarms

The categorization developed by Lees and Lee (2007) also depends on how drivers perceive different alarm types. Unlike other domains, driving involves a heterogeneous population with differing cognitive capabilities, skills, experiences, and personalities. Age is the most obvious way to distinguish drivers. Research suggests that alarm failures may have different consequences for different driver age groups (Dingus, et al., 1997; Fox & Boehm-Davis, 1998; NHTSA, 2005a). A recent field operational test found that compared to older drivers, younger and middle-aged drivers in the study rated the collision warning system more negatively and stated that FAs and NAs contributed to negative perceptions of the system (NHTSA, 2005a). Older drivers were more likely to select the higher sensitivity settings for the warning system. This suggests that older drivers may tolerate higher FA rates (Fox & Boehm-Davis, 1998).

Individual differences associated with age may influence the willingness to trust unreliable automation. Fox and Boehm-Davis (1998) investigated how system reliability influenced trust and compliance in an advanced traveller information system (ATIS) in younger (26-40) and older (66-80) drivers under four levels of reliability (40%, 60%, 80% and 100%). Overall, higher system reliability was associated with higher levels of trust. Older drivers trusted the system more than the younger drivers. These findings are somewhat contrary to another study that evaluated the effects of reliability (60%, 80%, 100%) on trust and reliance of a decision support aid in younger and older drivers (Sanchez, Fisk, & Rogers, 2004). Drivers were required to detect and count objects (e.g. pedestrians) while performing a monitoring task. The automation aided participants with the monitoring task but with varying reliability. Unlike younger drivers, older drivers differentiated systems with 60 and 80% reliability and expressed lower levels of trust in the 60% reliable system.

Differences in predisposition to trust or the levels or reliability examined may account for the differences between the two studies. For example, recent research has identified two constructs of trust that may be particularly important when examining how people view and interact with automation, dispositional trust and history-based trust (Merritt & Ilgen, 2008). Dispositional trust represents the level of trust upon initial interaction with automation and varies across individuals. In contrast, history-based trust (defined in this dissertation as system trust) represents the level of trust across interactions with automation and is a dynamic construct. Merritt and Ilgen (2008) found that people who have a higher pre-disposition to trust are more critical of “ill-functioning” automation. Consequently, the differences between the two studies may relate to the fact that reliability was a within-subjects variable for the Fox and Boehm (1998) study but a between-subjects variable in the Sanchez, Fisk and Rogers (2004) study.

Driver abilities and attitudes regarding ability may also be important when evaluating differences in how younger and older drivers tolerate and evaluate unreliable automation. Of particular interest here is how drivers 1) evaluate and identify potential hazards within a particular context, and 2) self-confidence – how drivers assess their ability to handle hazardous situations. These factors may influence how drivers characterize and respond to different alarm contexts (Zabyshny & Ragland, 2003).

Individual differences in hazard detection ability may influence system benefits and the perception of FAs

Hazard perception, the ability to recognize and anticipate hazardous roadway situations, is a critical driving related skill and correlates with accident involvement and driver experience (Horswill & McKenna, 2004; McKenna & Crick, 1994; Pelz & Krupat, 1974; Quimby & Watts, 1981). For example, Pelz and Krupat (1974) found that crash and conviction-free drivers responded 0.50 s quicker to hazards compared to drivers with a crash history and 1.20 s faster than drivers with violations (with or without a crash history). Horswill and McKenna (2004) found drivers with more crashes had worse hazard perception scores. Other researchers have found that novice drivers are less accurate and slower in perceiving hazards in filmed traffic situations compared to experienced drivers (Horswill & McKenna, 2004; McKenna & Crick, 1994). Hazard perception ability has also been shown to decline with age in drivers aged 65 and older (Horswill, et al., 2008). Cross-age studies indicate that the ability to perceive hazards peaks at age 55 and then subsequently declines (Quimby & Watts, 1981). These findings imply that both younger and older drivers have difficulty with hazard perception. However, different mechanisms likely contribute to diminished hazard perception ability and elevated crash risk within these groups.

Crash risk for younger drivers primarily relates to inexperience and the propensity to engage in risky behaviour (Deery, 1999; Mayhew, Simpson, & Pak, 2003; McGwin & Brown, 1999). Younger drivers possess acute perceptual and motor abilities, but may lack the experience needed to anticipate and avoid hazardous situations. Several researchers have argued that this reflects a lack of higher level cognitive skills, efficient attention allocation, and an overall awareness of one's surroundings that develop over time with experience (Fisher, Pollatsek, & Pradhan, 2006; Grayson & Sexton, 2002; Horswill & McKenna, 1999; Pollatsek, Fisher, & Pradhan, 2006). The propensity of these drivers to engage in risky behaviour may relate to an inability to appreciate risk. For example, Finn and Bragg (1986) examined risk perception using filmed traffic scenarios and found that younger drivers (aged 18 to 24) rated tailgating as less dangerous compared to middle-aged drivers (38 to 50 years old).

In contrast, crash risk for older drivers stems from age and disease-related declines in vision, cognition, and attention (Ball & Owsley, 1993; Owsley, Ball, Sloane, Roenker, & Bruni, 1991; Preusser, Williams, Ferguson, Ulmer, & Weinstein, 1998). For example, older drivers with useful field of view (UFOV) reductions show particularly elevated crash risk. UFOV represents the area from which visual information can be extracted during a single glance without making eye or head movements (Sanders, 1970). The measure incorporates visual sensory function, processing speed, divided attention, and selective attention (Ball & Owsley, 1993). UFOV reductions associated with age may co-occur with other impairments in vision, cognition, attention and memory. For example, UFOV may be sensitive to cognitive decline associated with mild cognitive impairment or Alzheimer's disease (Peterson, 2004; Rizzo, Anderson, Dawson, Myers, & Ball, 2000).

Although some have argued that UFOV impairments reflect a constriction in the in the field of view (Ball & Owsley, 1993), others have suggested alternative explanations or underlying features. For example, one study found that UFOV starts to deteriorate as early as 20 and reflects deficiencies in the ability to extract information from complex or cluttered scenes (Sekuler & Bennett, 2000). Recent work suggests that UFOV may also reflect an inability to disengage attention from a previously attended location (Cosman, Lees, Vecera, Lee, & Rizzo, In Revision). While the mechanisms related to UFOV are largely in question, it does appear that UFOV represents a useful tool for identifying at-risk drivers.

Several studies have shown that UFOV scores are associated with driving performance outcomes derived from state crash records, on-road driving tests, and driving simulators (Ball & Owsley, 1993; Goode, et al., 1998; Myers, Ball, Kalina, Roth, & Goode, 2000; Roenker, Cissell, Ball, Wadley, & Edwards, 2003). These studies suggest that drivers with attention impairments make more driving errors and have higher crash rates than drivers without impairments. One study found that reductions in hazard perception ability coincided with diminished contrast sensitivity, larger reductions in the UFOV and slower simple reaction time performance (Horswill, et al., 2008).

At the same time, older drivers have extensive experience that may help them to detect potential conflicts and to accommodate for such deficiencies. Rear-end crashes for these drivers may occur because they require more time to process information, make a decision, and implement a response (Lerner, 1994). Some research suggests that older drivers may attempt to compensate for limitations by reducing their exposure, avoiding risky situations, and modulating their driving style (e.g., adopting slower speeds) (Ball, et al., 1998; Charlton, Oxley, Fildes, & Les, 2001). Even when such measures are taken they may not be sufficient enough to reduce the risk of crashing (Ball, et al., 1998). As well, some older drivers, such as those with UFOV impairments, may fail to adequately compensate if they are unaware of their functional impairments (Charlton, et al., 2001; Owsley, et al., 1991).

The different mechanisms underlying the elevated crash risk of younger and older drivers suggests that collision warning systems may aid both groups, but through different mechanisms. Novice drivers can improve their hazard perception abilities through repeated exposure to hazards or by receiving training on how to identify hazards (Fisher, et al., 2006; Grayson & Sexton, 2002; McKenna & Crick, 1994; McKenna, Horswill, & Alexander, 2006). For instance, McKenna, Horswill and Alexander (2006) demonstrated in a series of experiments that hazard detection and speed modulation could be improved using commentated training videos. For younger drivers' collision warning systems may represent a means to train drivers to identify and anticipate hazards. These systems may encourage younger drivers to modify their behavior (e.g., by promoting drivers to adopt safer headways) such that drivers reduce their exposure to risk. In contrast, for older drivers the primary benefit might be supplementing bottom-up processing and decision making. Specifically, these systems direct attention to objects that the driver might otherwise overlook or be slow to respond to, thereby allowing drivers more time to process the situation and determine an appropriate response.

When operating in parallel with automation (i.e., collision warning system), the driver becomes responsible for filtering information the system provides and either agrees or disagrees with the diagnosis given based on their own hazard assessment (Bisantz & Pritchett, 2003). How individuals assess hazards influences the effect and frequency of subjective alarm failures. Drivers are likely to have a greater appreciation for warnings that occur in situations they find hazardous. Therefore, the same factors that increase crash risk for younger and older drivers may prevent them from recognizing hazards and the value of warnings (Lees & Lee, 2009), and in the worse case scenario, might lead younger and older drivers failing to appreciate TAs.

Likewise, responses to alarms may not only depend on the driving context but also the driver's capabilities. For example, even if NAs occur consistently in similar situations, older drivers may have more difficulty identifying what triggered the alarm. Older drivers with UFOV impairments may have greater difficulty identifying the source of an alarm that occurs in a complex traffic situation. Cognitive declines in this group may also impede them from understanding why the system has generated a warning. In contrast, younger drivers may not fully appreciate UAs and TAs because they might not consider the situations associated with these warnings as hazardous.

A major goal of this dissertation is to determine if different driver groups make similar hazard assessments. This will help identify what situations create discrepancies between the driver and the system and whether such discrepancies differ by age. Individual characteristics associated with age may influence the drivers' agreement with the system and the drivers' ability to understand how the system is operating. One would expect greater system distrust when the drivers' hazard assessment is incongruent with the system and when the driver is unable to understand what triggered the alarm.

Individual differences in self-confidence may influence system benefits and the perception of FAs

Self-confidence represents another important construct that may influence the benefits of collision warning systems and the perception of alarms. In terms of driving, self-confidence refers to the perception that a driver has the ability or skills to handle hazardous situations. A variety of techniques (e.g., questionnaires, traffic scene evaluation, filmed traffic scene evaluation) provide convergent evidence of systematic bias that influences how people, particularly younger and older drivers, will respond to collision warnings.

Drivers generally tend to consider themselves as more skillful, more safe, and less likely to crash compared to their peers and in some cases even other age groups (DeJoy, 1992; Finn & Bragg, 1986; Horswill, Waylen, & Tofield, 2004; Matthews & Moran, 1986; Waylen, Horswill, Alexander, & McKenna, 2004). Such research has primarily focused on younger/novice drivers, finding that younger inexperienced drivers, especially males, are more likely to overestimate their driving skill and perceive lower levels of risk than middle-aged/experienced drivers (Deery, 1999; DeJoy, 1992; Finn & Bragg, 1986; McKenna, Stanier, & Lewis, 1991; Tränkle, Gelau, & Metker, 1990). Some of these studies have emphasized the role of driving experience, finding that age and sex differences diminish when driving experience is taken into account (Groeger & Brown, 1989; McKenna, et al., 1991).

Like hazard detection, crash risk and skill estimation, depend not only the driver making the estimation but the type of situation or skill being assessed (Finn & Bragg, 1986; Horswill, et al., 2004; Matthews & Moran, 1986). For example, one study examined biases in both vehicle control and hazard perception skill using subjective ratings and found that biases were greater for hazard perception skills compared to vehicle control and overall driving skill (Horswill, et al., 2004). Another study found that drivers who have more self-confidence may view situations as less risky (Matthews & Moran, 1986).

Self-confidence has also been shown to influence how people use automation (Gao & Lee, 2006; G. Ho, Wheatley, & Scialfa, 2005; Kantowitz, et al., 1997; Lee & Moray, 1994). Lee and Moray (1994) found that both trust and self-confidence determined how much operators relied on automation to control a semi-automatic pasteurization plant. When trust in the automation exceeded self-confidence people used automatic control. In contrast, when self-confidence exceeded trust people avoided using the automation and instead performed the task manually. A study conducted to determine whether age influenced trust and reliance on a medication management system found that compared to older adults, younger adults had more self-confidence in their abilities and less trust in the automation (G. Ho, et al., 2005). The attitudes adopted by the different age groups were associated with different patterns of automation use. Specifically, younger participants relied on the automation less. Another study found that trust and reliance declined when an IVIS system experienced errors in a familiar setting where the driver was more confident about making route decisions than in an unfamiliar setting where self-confidence was lower (Kantowitz, et al., 1997). These findings suggest that drivers who are more confident in their abilities and less trusting of automation may rely on automation less.

Research suggests that even though novice drivers have the highest crash risk they may overestimate their driving abilities and underestimate their crash risk (Gregersen, 1996). Consequently, while these drivers may benefit most from collision warning systems they may be less likely to comply with warnings. Confident drivers may also be more critical of systems that generate certain types of alarms. In contrast, older drivers may be less confident regarding their ability to carry out certain manoeuvres and may trust automation more. The tendency of older drivers to adapt their exposure and avoid certain manoeuvres provides some evidence that these drivers may be less confident. Combined these conditions may lead older drivers to be less critical of unreliable automation compared to younger and middle aged drivers.

Individuals' driving style influences system reliability

The FA rate and consequently the value of a warning system can vary as a function of the drivers actions (Meyer & Bitan, 2002). When a driver takes action to minimize negative outcomes they lower the PPV and increase the negative predictive value of a system. Therefore, the better a driver is at avoiding negative outcomes, the less likely the system will be beneficial.

Driver behaviour influences system performance such that the type and frequency of failures often varies across drivers (Ben-Yaacov, et al., 2002; Dingus, et al., 1997; Maltz & Shinar, 2004; NHTSA, 2005a). In a field operational test, the number of imminent alarms issued to drivers varied considerably with a range of 0.08/100 miles to 4.34/100 miles (Algorithm C, NHTSA, 2005a). In on-road and naturalistic studies, the type of failure (FA or missed event) depends on the driving style of the driver (Ben-Yaacov, et al., 2002; Dingus, et al., 1997; Maltz & Shinar, 2004; Sullivan, et al., 2008). Drivers who adopt longer headways will receive more FAs because no hazard exists under these conditions. Conversely, drivers who adopt shorter headways experience more missed events but they will also receive more TAs.

Differences in the driving style of younger, middle-aged, and older drivers may result in these groups receiving different types of failures (Meyer & Bitan, 2002; NHTSA, 2005a). For example, younger drivers may receive more FAs because they are more willing to adopt shorter headways (Evans & Wasielewski, 1982, 1983). Sullivan, Tsimhoni, and Bogard (2008) evaluated how the reliability of a lateral drift warning system influenced driving and found that the number of alarms received by younger, middle-aged and older drivers differed substantially. The authors had to exclude data from several older drivers because they received far fewer alarms than drivers in the other age groups.

Differences in the number and types of alarms received by drivers may make it difficult to understand what changes behaviour when such systems are present (Dingus, et al., 1997; NHTSA, 2005a). Dingus et al. (1997) investigated how FAs influence the benefits of a visual/auditory headway maintenance and collision warning system in younger (18 to 24 years old) and older (65 and older) drivers. Drivers completed a baseline condition and one of four FA rate conditions. The number of alarms was highly variable and consequently the original FA groups were re-categorized as 0-30%, 31-60% and >60%. For coupled headway events, older drivers adopted safer following distances compared to younger drivers. FAs caused younger drivers to adopt larger minimum headways. However, the benefits of the system eroded for younger drivers when the FA rate exceeded 60%. Older drivers adopted a long, relatively fixed headway time independent of whether the system was present or the number of FAs received. That is, the FA rate did not influence the following behaviour of older drivers.

Dissertation objectives

Rear end collisions occur frequently and collision warning systems might reduce their occurrence. However, the benefits of these systems depend on driver trust and compliance, both of which can be undermined by alarm failures. The FA problem is more complicated than defining a threshold of reliability in which a system becomes unacceptable to the user. A number of factors might influence whether or not drivers trust and comply with imperfect collision warning systems.

The goal of this dissertation is to isolate how and why different alarms influence drivers and whether this influence depends on age. Figure 8 provides a theoretical framework describing the primary factors assumed to influence drivers' perception and response to different alarm contexts. This model presents the main components thought to influence alarm perception examined in this dissertation. Two aims are addressed in this dissertation and are outlined in Figure 8 with components in grey examined under aim 1 and components in black and grey examined under aim 2.

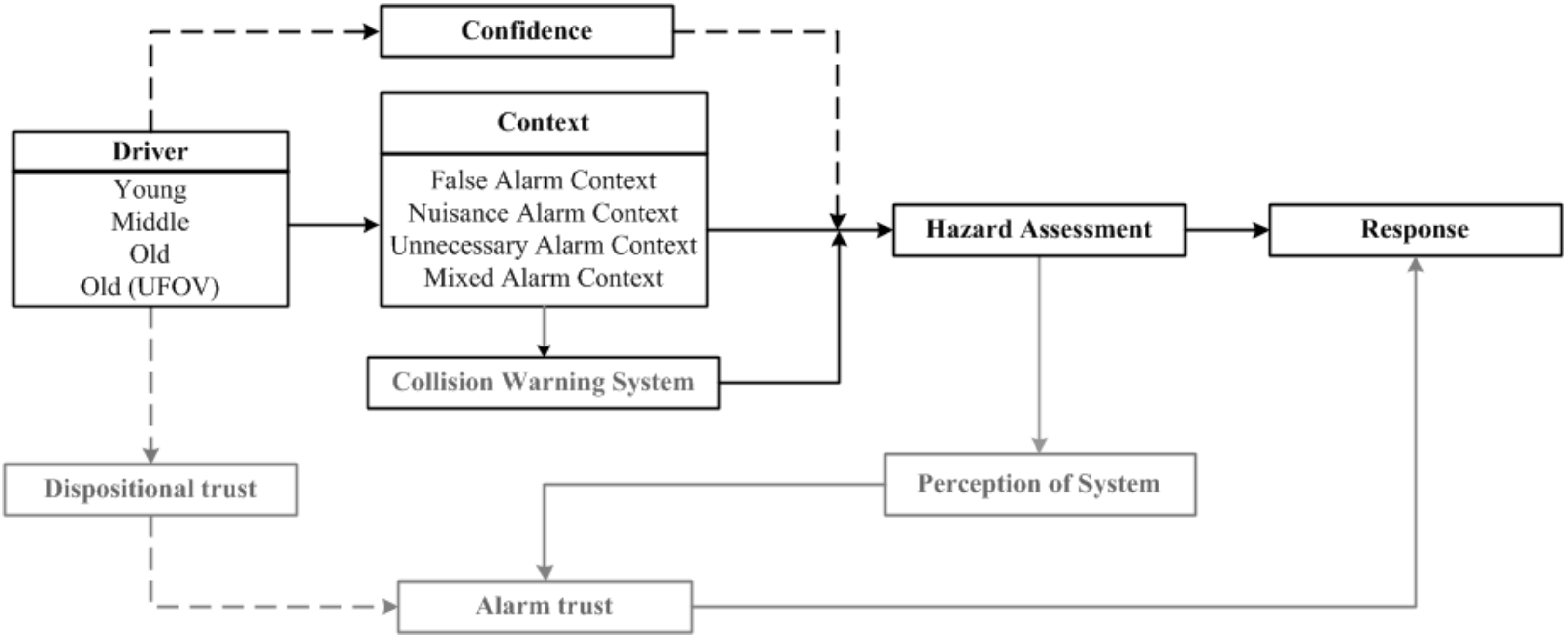

Figure 8. Overview of factors expected to explain how different alarm types influence trust and compliance.

Aim 1: Determine how self-confidence and hazard perception differ between younger, middle-aged, and older drivers.

Drivers exposed to different driving contexts may or may not respond to the situation they are confronted given their evaluation of hazards (i.e., hazard assessment). The driver's self-confidence – how drivers assess their ability to handle situations – represents one factor that can influence hazard assessment.

Driving context is expected to influence hazard assessment and how drivers respond. Specifically, drivers will perceive FA and NA contexts as less hazardous than UA and TA contexts. Diminished hazard perception in younger and older drivers with attention impairments will cause these groups to view UA and TA contexts as less hazardous than middle-aged and older drivers. As a result younger and older drivers with UFOV impairments will respond less frequently during these events.

Drivers will also differ in how confident they are when faced with different situations. Younger drivers are expected to be more confident in their ability to avoid a collision compared to other age groups. Although research on older drivers is somewhat limited with respect to these measures, it is expected that older adults (especially those with UFOV impairments) will be less confident.

Aim 2: Assess how younger, middle-aged, and older drivers perceive different types of alarms.

Collision warning systems generate alarms within different driving contexts. Based on the alarm and the context the driver assesses hazards and determines whether to respond. System perceptions are influenced by the alarm context and whether the driver agrees with the systems hazard assessment. In addition, some drivers are more trusting than others (dispositional trust) and consequently may be more trusting of unreliable automation.

It is expected that different alarms will differentially influence alarm perception, and trust. UAs and TAs are more consistent with drivers hazard assessment and therefore drivers will trust these alarms more than NAs or FAs. At the same time, the predictability of NAs will lead drivers to trust them more than FAs. The effects of different alarms are expected to be influenced by age.

In general, differences in hazard assessment and self-confidence will influence how drivers of different age groups trust different alarms. Younger drivers are expected to trust alarms less than middle-aged adults, regardless of the alarm type experienced. Older drivers are expected to trust alarms more compared to middle-aged and younger drivers. In addition, for older drivers who are expected to rate situations as more hazardous and to be less confident in their abilities alarms will increase the drivers hazard assessment compared to when no alarm occurs.

CHAPTER 3.

METHODS

This dissertation uses a subset of videos collected during the Automotive Collision Avoidance Field Operational Test (ACAS FOT) (NHTSA, 2005a) to examine how younger, middle-aged, and older – with and without useful field of view (UFOV) impairments— drivers assess different alarm contexts with and without collision warning system warnings. The overall goal is to determine how different types of alarms influence trust and hazard evaluation.

Participants

Previous studies that have used a similar task setup suggest that a sample size of 12 to 16 per group is generally necessary and sufficient for finding meaningful differences between different groups (Graham, 1999; Horswill, et al., 2008; McKenna & Crick, 1994). One hundred and twenty five drivers completed the study, 1) 32 younger drivers (M: 19.2, S.D.: 0.7, 16 female and 16 male), 2) 32 middle-aged drivers (M: 43.4, S.D.: 4.1, 16 female and 16 male) 3) 32 older drivers without UFOV impairments (M: 72.9, S.D.: 5.8, 15 female and 17 male), and 4) 29 older drivers with UFOV impairments (M: 77.3, S.D.: 4.8, 13 female and 16 male). Half the participants within each group watched the videos with auditory warnings and half without.

Participant Recruitment

Younger, middle-aged, and older participants were recruited from the Iowa City and Coralville area through newspaper ads, flyers, and university wide mass emails. The majority of older drivers with and without UFOV impairments participated as part of a larger examination of these drivers.

All drivers were screened for UFOV impairments using the Visual Attention Analyzer (Model 3000, Vision Resources, Chicago, IL, Ball & Owsley, 1993; Edwards, et al., 2005) which measures performance (the display duration (ms) required to attain 75% response accuracy) on four subtests. Subtest 1 (processing speed) measures performance during a two-alternative forced choice task presented in central fixation. Subtest 2 (divided attention) measures how fast subjects concurrently identify central and peripheral targets. Subtest 3 (selective attention) resembles subtest 2 but the peripheral target is surrounded by distracters. Subtest 4 is similar to subtest 3 but the central task requires a same/different discrimination.

Performance on subtests 3 and 4 were used to identify older drivers with UFOV impairments. Older drivers with a score greater than or equal to 350 on subtest 3 and/or 500 on subtest 4 were classified as having impairments. This criterion was used because it corresponds to previously used cut-offs (Edwards, et al., 2005) that had a sensitivity of 89% and specificity of 81% for predicting crash involvement (Ball & Owsley, 1993). For example, in one study older drivers who failed the test had about 4.2 times more accidents than older drivers who passed (Owsley, et al., 1991).

Table 2. Mean (standard deviation) for each UFOV subtest

	Subtest 1 Processing Speed	Subtest 2 Divided Attention	Subtest 3 Selective Attention	Subtest 4 Same/Different
Younger	16.13 (0.71)	17.81 (5.98)	60.994 (37.35)	173.75 (89.02)
Middle	16.47 (1.54)	22.00 (13.94)	69.41 (25.58)	179.66 (83.95)
Older	19.09 (4.80)	47.91 (46.15)	180.22 (67.43)	320.03 (88.89)
Older (UFOV impaired)	25.07 (14.36)	63.71 (56.35)	354.38 (153.11)	500 (0)

Participant screening

A phone screen was used to determine eligibility to participate in the study (Appendix A). Drivers taking certain medications or with existing medical conditions (e.g. neurodegenerative disease, anxiety, depression) that might influence performance were excluded from participating in the study. All participants were required to have an active driver's license, to report having normal to corrected normal vision, and to be a native English speaker. Younger drivers were required to drive 100 miles/week or less and middle-aged drivers were required to drive 100 miles/week or more. Middle-aged drivers were screened on their violation and crash history. If the number of crashes and violations divided by the number of years driving was greater than 0.50 the driver was excluded from participating in the study.

Table 3. Descriptive data for younger, middle-aged, older and older(UFOV) drivers that participated in the study.

	N	Age	Years Driving	Miles per Week	Crashes	Moving Violations	Near Visual Acuity	Far Visual Acuity	Contrast Sensitivity
Younger	32	19.22 (0.71)	3.97 (1.12)	26.56 (29.85)	0.44 (0.62)	0.28 (0.52)	0.01 (0.01)	-0.02 (0.08)	1.80 (0.16)
Middle	32	43.38 (4.13)	27.22 (4.47)	298.75 (274.95)	0.16 (0.45)	0.28 (0.52)	0.04 (0.07)	-0.03 (0.07)	1.88 (0.15)
Older	32	72.88 (5.77)	55.61 (6.14)	125.00 (99.79)	0.19 (0.40)	0.13 (0.34)	0.04 (0.04)	0.12 (0.12)	1.70 (0.17)
Older (UFOV)	29	77.34 (4.77)	57.52 (8.76)	111.24 (120.43)	0.21 (0.41)	0.10 (0.31)	0.09 (0.08)	0.19 (0.16)	1.54 (0.21)

Apparatus

An Intel® Pentium® D processor-based PC displayed the videos on a 20 inch monitor and recorded responses and response times for the hazard perception task. Another PC with the same monitor size displayed and collected questionnaire data (developed and administered using surveygizmo). The participant was situated approximately 2 feet from the monitors. The overall setup is presented in Figure 9.

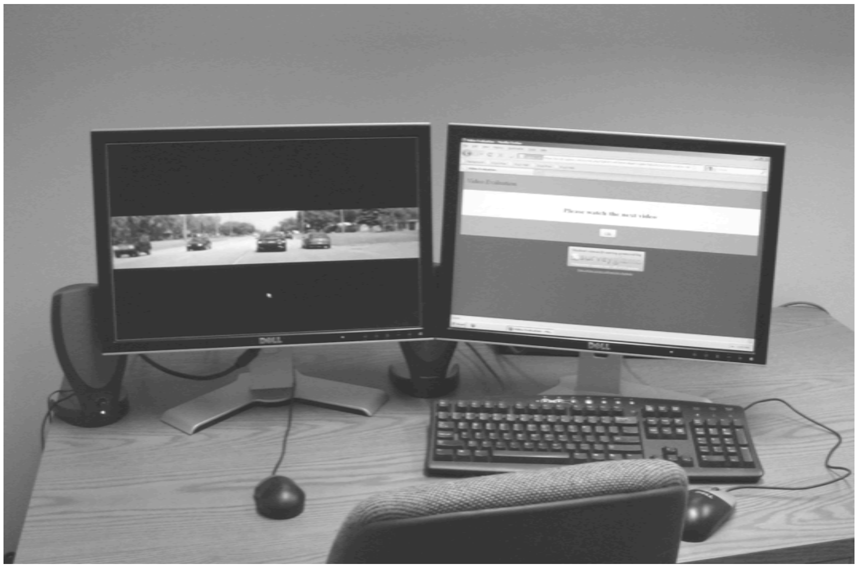

Figure 9. Experimental setup. The monitor on the left was used to play the videos and record responses. The computer on the right was used to administer the questionnaires. Auditory warnings, when used, were played using standard PC speakers located beside the monitor on the left.

Hazard Perception Test

Researchers have examined reliability issues with a variety of methods inspired by different perspectives and research objectives. Such methods include the use of surrogate experiments, and simulation. This dissertation used a hazard perception task to evaluate how drivers assess and respond to different driving contexts with and without a collision warning system alarms. A similar approach has been used by other researchers to examine hazard assessment (Finn & Bragg, 1986; Matthews & Moran, 1986), different auditory warnings (Graham, 1999), and the influence of in-vehicle train warnings on trust and perception response times (Chugh & Caird, 1999).

There are several benefits and limitations to using this approach. First, this approach allows a high level of controllability in terms of how many alarms (either true or false) drivers receive and the conditions in which they are triggered. Using this approach every driver receives the same conditions and same failure rate. Second, because the videos incorporated were collected from the ACAS FOT that used a forward collision warning system (NHTSA, 2005a), the types and timing of alarms generalize to those experienced by real-world systems. Third, this method allows a high level of visual realism and complexity that other methods simplify. At the same time, a limitation of this approach is that the driver is unable to control the conditions they are exposed to.

Hazard Types

The number of clips used for such tasks varies quite substantially often ranging from 20 to 40 clips. Some researchers have used as few as eight clips to evaluate hazard perception (McKenna, et al., 2006). In this experiment, four separate hazard perception tasks, with ten videos each, were created using 40 video files collected as part of the ACAS FOT as provided by UMTRI (NHTSA, 2005a). The original video data was collected at 10 Hz for 5 seconds before and 3 seconds after the system issued an imminent alarm. Over 1500 events (audible or silent alarms) occurred during the field operation test in which 96 drivers interacted with a collision warning system.

The 40 videos were equally distributed across four tests. Those selected were chosen based on the framework provided by Lees and Lee (2007). For each hazard perception test, six videos were selected to represent a particular type of alarm failure and four videos were selected to represent a true alarm (TA) context. A reliability of 40% was chosen to make it more likely that these systems would be prone to inducing diminished trust and compliance. In three of the four tests, only one type of failure was examined. In the fourth a mix of alarm failures was examined.

Failure Context

The videos selected were chosen based on the target that instigated the imminent collision warning and the context surrounding the alarm onset. The videos were divided into four different alarm types, defined in table 4, using the theoretical framework provided by Lees and Lee (2007): 1) false alarm (FA), 2) nuisance alarm (NA), 3) unnecessary alarm (UA), and 4) true alarm (TA).

Table 4. Naming scheme used to differentiate alarm contexts

Alarm type (warning)	**Context** (warning & no warning)	**General Definition**
False alarm (FA)	FA Context	An alarm associated with a non-hazardous context where the operator is unable to identify the source (e.g., system malfunction) and a failure to respond is not associated with a poor outcome.
Nuisance alarm (NA)	NA Context	An alarm associated with a non-hazardous context where the operator can identify the source but a failure to respond is not associated with a poor outcome.
Unnecessary alarm (UA)	UA Context	An alarm associated with a potentially hazardous situation but where the situation resolves itself such that a failure to respond is not associated with a poor outcome (e.g., a collision)
True alarm (TA)	TA Context	An alarm associated with a hazardous situation where the operator must intervene to avoid a poor outcome (e.g., a collision)

Examples of each alarm type are presented below. In each image, the numbers correspond to the stationary and moving targets being tracked by the radar. Circles identify the target that initiated the imminent warning. It should be noted that in all of the videos used the system determined a hazard was present and that the driver should respond.

FA context (Figure 10): cases in which the system was triggered by an out-of-path obstacle. All videos were selected with the assumption that it would be difficult for the driver to identify what caused the alarm. For example, in the videos chosen there were no vehicles in front of the driver, and the objects triggering the alarms were quite variable. The goal was to make the alarms associated with these situations appear random. For example, in Figure 10 the target (identified by a circle) that instigated the alarm was 1) a tree or small post, and 2) an oncoming vehicle in the adjacent lane. A video review of the ACAS FOT (NHTSA, 2005A). A video review of the FOT data was conducted and verified that for the videos used in this experiment the driver did not respond.

Figure 10. Example FA context. The video used was collected as part of the ACAS FOT study (NHTSA, 2005A).

NA context (Figure 11): cases in which the system was triggered by an out-of-path obstacle. Similar to FA contexts, these contexts involve the system being triggered by an out-of-path obstacle. However, drivers may understand the source of the alarm within these contexts. Two scenarios were incorporated and are depicted in Figure 11: a) a roadside construction barrel, or b) signage while coming around a corner. A video review of the FOT data was conducted and verified that for the videos used in this experiment the driver did not respond.

Figure 11. Example NA context. The video used was collected as part of the ACAS FOT study (NHTSA, 2005A).

UA context (Figure 12): cases where the system was triggered by a transitioning-host path scenario where the conflict vehicle made a lateral movement that resolved the conflict. Two types of transitioning-host path scenarios were incorporated and are depicted in Figure 12: a) a vehicle turned across the driver's path, or b) a vehicle slowed down up ahead to make a turn. The driver's hazard assessment of these situations may more closely match that of the system. A video review of the FOT data was conducted and verified that for the videos used in this experiment the driver did not respond.

Figure 12. Example UA context. The video used was collected as part of the ACAS FOT study (NHTSA, 2005A).

*TA context (Figure 12)***:** cases where the system was triggered because the situation required the driver to respond to avoid crash involvement. As shown in Figure 13, these are in-host path scenarios where the lead vehicle (LV) is the target that instigated the alarm. In these situations, the driver's assessment of the situation should more closely match the system, which has determined a hazard is present and the driver should respond. These clips were identified in a conflict analysis conducted by VOLPE (Najm, et al., 2006) that determined 28 of the alarms occurring during the ACAS FOT (NHTSA, 2005A) were TAs when algorithm C was used. A video review of the FOT data was conducted and verified that for the videos used in this experiment the driver did respond. Videos in which a driver’s response was apparent (e.g. hard braking caused the video to shift, or swerving) were excluded.

Figure 13. Example TA context. The video used was collected as part of the ACAS FOT study (NHTSA, 2005A). Videos were identified in a conflict analysis conducted by VOLPE (Najm, et al., 2006).

Auditory Warnings

To determine how drivers evaluate these contexts with these systems, half the participants watched the videos with the alarms muted and half the participants received alarms while completing the hazard perception task. The ambient sound level of the room was 56 dBA. Based on recommendations that auditory alarms should be 10 to 15 dBA above the ambient sound level, auditory alarms were played at 70 dBA through standard PC speakers (Graham, 1999). Other studies have used similar sound levels when investigating such systems (Bliss & Acton, 2003; Lee, et al., 2002; Lees & Lee, 2007). The auditory alarm, used to inform the driver of impending conflicts, consisted of an abstract auditory alarm used in previous studies that was shortened from 2.25 s to 200 ms (Lee, et al., 2002; Lees, et al., 2009; Lees & Lee, 2007; Tan & Lerner, 1995). For all videos, the timing of the warning was based on the actual timing (of imminent alarms) that occurred during the ACAS FOT (NHTSA, 2005A).

Experimental design and independent variables

As shown in Figure 14 the experiment is a 4 x 4 x 2 mixed design with alarm type blocked (FA, NA, UA, and mix of “false” alarms) as a within-subjects factor, and driver group (younger, middle, old, and old with UFOV) and warning (present, absent) as between-subjects factor.

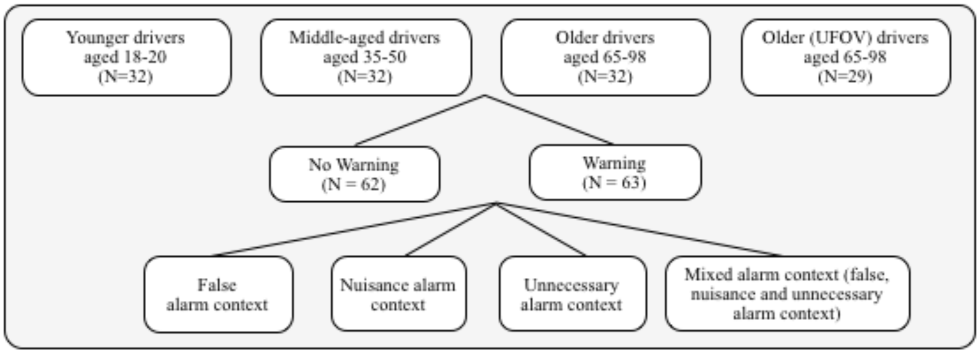

Figure 14. Experimental design

Procedure

Upon arrival, participants read and signed an informed consent and the researcher addressed any pending questions. Participants completed a demographic questionnaire (Appendix B), visual acuity and contrast sensitivity vision tests, and the UFOV task. Participants also completed an interpersonal trust (Appendix C) and driver skills questionnaire (Appendix D) using a PC.

For the main task, participants were told that they would be watching a set of videos and that after each video they would be asked to complete a set of questions. They were given an instruction sheet explaining the videos and how they should respond during the task (Appendix E). Prior to the practice, the researcher reiterated the following points: 1) participants should pretend they were driving the vehicle and that the videos corresponded to what they were seeing out the front window, 2) the participants task was to determine if they would respond in the situation, 3) if participants felt the situation required a response they should press the mouse at the time they, as the driver, would initiate a response. Half of the participants were informed that they would be evaluating the videos with different collision warning systems. Their instruction sheet indicated that they would be receiving auditory alarms from a forward collision warning system to help them identify traffic conflicts. The instruction sheet indicated that they would be interacting with four different systems. Another sheet was used to explain the general purpose of the forward collision warning system (Appendix G).

Participants were given five practice trials to become familiar with the task and to gain experience with completing questionnaires in between video presentation. All of the practice trials involved a situation where a LV braked ahead of the host vehicle and were situations where the ACAS FOT driver had made a response (NHTSA, 2005A). After each video drivers completed a set of questions (Appendix F for drivers who completed the task without warnings and Appendix H for those who completed the task with warnings. Subsequently drivers completed four experimental blocks that were counterbalanced using a Latin square. Sixteen different video orders were created for each block. After completing the experimental task, participants were debriefed. Sessions normally lasted 1 to 2 hours for the no warning condition and 1.5 to 2.4 hours for the warning condition. Participants were compensated $10 per hour.

Outcome measures

Three groups of outcome measures were examined across the warning and no warning conditions: 1) self-confidence, 2) hazard assessment, and 3) driver responses.

1. *Self-confidence*. After each video, drivers were asked how confident they were that they could avoid a collision in situations such as the video just viewed (on a scale of 0 to 4, 0 being not at all confident and 4 being extremely confident).

2. *Hazard assessment*. Drivers were asked to indicate how hazardous the situation they watched was (on a scale of 0 to 4, 0 being not at all hazardous and 4 being extremely hazardous).
3. *Driver response:* represents whether the driver responded to the situation and the time of the response in relation to the time a warning did or would have occurred.

For drivers who received warnings during the experiment, alarm trust and alarm perceptions were measured using a five-point rating scale.

1. *Alarm trust*. After each video, drivers were asked to indicate the trustworthiness of the system. Responses were made using a scale of 0 to 4, 0 being not at all trustworthy and 4 being extremely trustworthy.
2. *Alarm perceptions*. After each video, drivers were asked to rate how understandable, useful and appropriate the alarm was within the current context. Responses were made on a scale of 0 to 4, 0 being not at all and 4 indicating extremely.

Two aims will be addressed with the data collected: 1) determine how self-confidence and hazard perception differ between younger, middle-aged, and older drivers and 2) assess how younger, middle-aged, and older drivers perceive different types of alarms.

Chapter 4 addresses aim 1 by determining how younger, middle-aged and older drivers differ in terms of self-confidence and in how they assess and respond to different alarm contexts. Three outcome measures are examined across the warning and no warning conditions: 1) confidence, 2) hazard assessment, and 3) driver responses.

Figure 15 shows how hazard ratings might differ for the different alarm contexts and age groups. Responses are expected to follow a similar pattern. Drivers are expected to be extremely confident during FA and NA contexts because drivers will not view these situations as hazardous. As a result, drivers will rarely respond while viewing these situations.

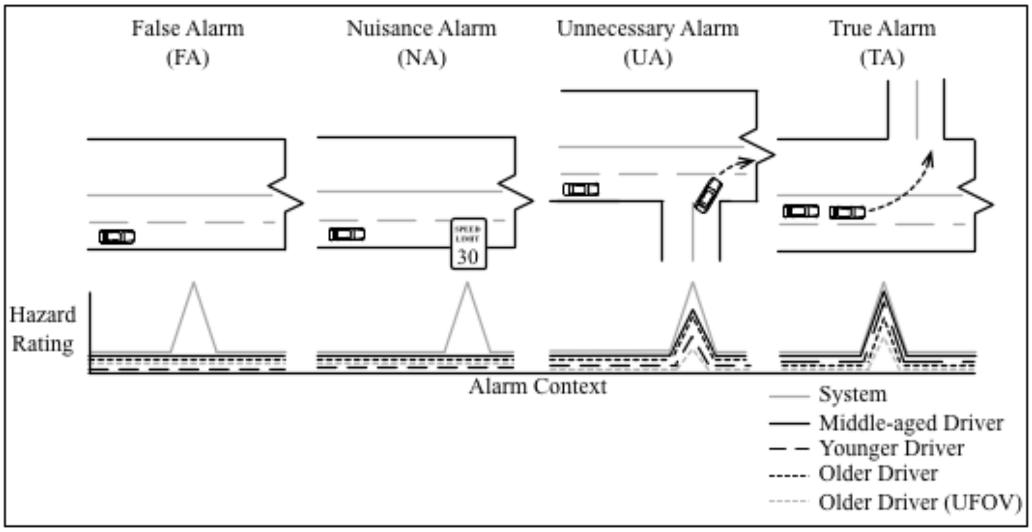

Figure 15. Hazard ratings for younger, middle-aged, older drivers (with and without UFOV impairments) and a collision warning system for FA, NA, UA, and TA contexts.

A very different pattern of results is expected for UA and TA contexts. It is expected that, regardless of whether warnings are provided, drivers will consider these situations as more hazardous than FA and NA contexts. Compared to middle-aged and older drivers, younger drivers are expected to view these situations as less hazardous, to brake less frequently and to be more confident. In contrast, older drivers are expected to indicate that these situations are more hazardous than younger drivers. Older drivers with UFOV impairments are expected be view these situations as less hazardous and brake less frequently than older drivers without impairment. Overall, older drivers with and without impairments will be less confident than the other driver groups.

It is expected that drivers will respond to the different alarm contexts in a similar manner across the warning and no warning condition. The one exception is that alarms are expected to increase hazard ratings and response frequency for UA contexts especially for older drivers who may be less confident in their abilities.

Chapter 5 addresses aim 2 and examines how younger, middle-aged, and older drivers perceive alarms that occur within different alarm contexts. Overall, it is expected that older drivers will be more trusting of collision warning systems compared to middle-aged or younger drivers. It is also expected that differences in how drivers perceive alarms that occur within different driving contexts will influence trust. FAs will result in a high degree of distrust because drivers will not consider these contexts hazardous and it is difficult for drivers to determine what caused the alarm. In contrast, drivers will identify what caused NA, UA, and TA to occur. A greater understanding of how the system is operating will lead drivers to trust NAs more than FAs. Trust is expected to be highest for the UAs where failures may be both understandable and useful. It is expected that because drivers will consider UAs and TAs as more useful and understandable than other alarm types, in part because drivers view these situations as more hazardous.

CHAPTER 4.

THE INFLUENCE OF DRIVER GROUP AND WARNINGS ON HAZARD ASSESSMENT IN DIFFERENT DRIVING CONTEXTS

This chapter examines differences between younger, middle-aged and older drivers in relation to self-confidence, hazard assessment, and responses when exposed to different alarm contexts. Data was analyzed using a mixed model in SAS (Statistical Analysis System) version 9.2. Group and warning were included in the models as between-subjects variables. Alarm context (FA, NA, UA, and TA context) was treated as a within-subjects variable repeated across each block. Subject was treated as a random effect and the compound symmetry covariance structure was used based on the ability of the model to converge and AIC values. Pairwise comparisons were computed with Tukey-Kramer adjustments. Table 5 provides an overview of the main and interaction effects for confidence, hazard ratings, and response frequency. These findings are further discussed throughout the chapter.

Table 5. Overview of statistical main and interaction effects for confidence ratings, hazard ratings, and response frequency.

			F Value		
Effect	Num DF	Den DF	Confidence Rating	Hazard Rating	Response Frequency
Group	3	118	5.23**	7.72***	0.71
Warning	1	118	0.21$^{\pm}$	2.40	0.98
Alarm Context	3	351	16.89***	533.06***	1791.30***
Group x Warning	3	118	2.32***	1.33	1.63
Group x Alarm Context	9	351	11.33***	7.81***	19.64***
Warning x Alarm Context	3	351	46.32***	13.78***	1.49
Group x Warning x Alarm Context	9	351	7.00***	2.55**	3.00**

±$p \leq 0.10$, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

Driver confidence

As shown by table 5, the results partially support the expected outcomes for confidence across the different driver groups and alarm contexts. As anticipated younger drivers (M = 3.80, S.E. = 0.01) reported being more confident then older drivers with and without UFOV impairments (older: M = 3.44, S.E. = 0.02, older UFOV: M= 3.33, S.E. = 0.02; table 6). Surprisingly, the middle-aged group (M = 3.59, S.E. = 0.01) was not significantly more confident then either of the older driver groups.

Table 6. Significant pairwise comparisons for main effects for confidence ratings.

	Estimate (Δ)	t-value	d.f.	95% CI (Adj.)
Group				
Younger vs. Older	-0.352	2.88^{*}	118	(-.671, -0.034)
Younger vs. Older (UFOV)	-0.472	3.77^{**}	118	(-0.799, -0.146)
Alarm Context				
FA vs, NA	-0.089	-3.41^{**}	351	(-0.156, -0.022)
FA vs. TA	0.0648	2.87^{*}	351	(0.006, 0.123)
NA vs. UA	0.1406	5.38^{***}	351	(0.073, 0,208)
NA vs TA	0.1237	6.80^{***}	351	(0.100, 0.212)

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

A somewhat surprising pattern emerges when confidence is evaluated in terms of group, warning, and alarm context. Figure 16 plots mean (±1 S. E.) confidence across the different driver groups and different alarm contexts for the no warning (top) and warning conditions (bottom). As expected, younger drivers reported consistently high levels of confidence for the warning and no warning conditions for all of the alarm contexts.

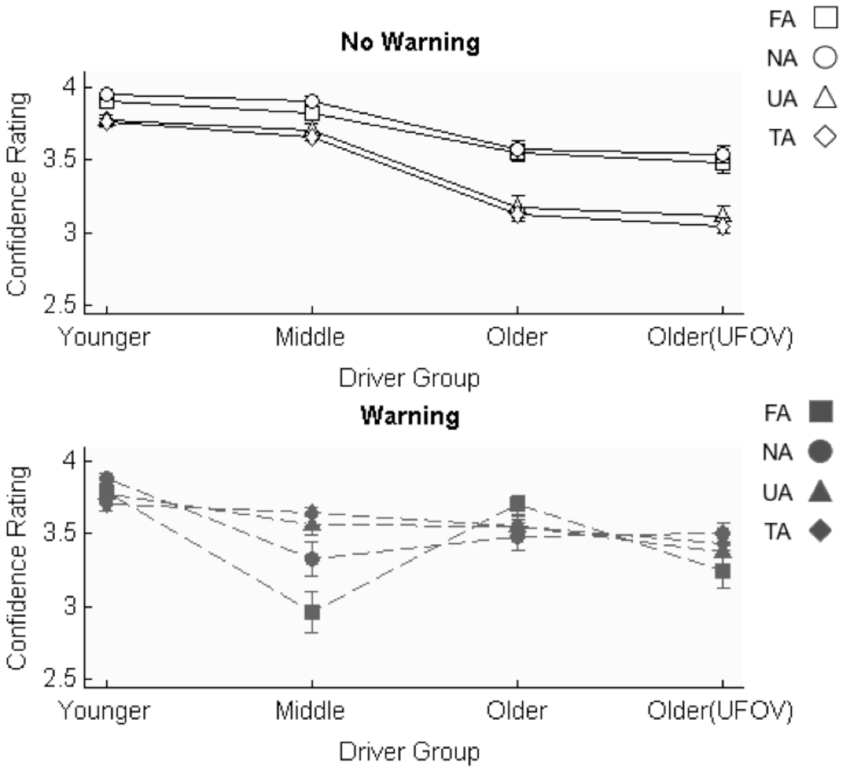

Figure 16. Mean (±1 S. E.) confidence ratings for younger, middle-aged, and older drivers (with and without UFOV impairments) for the warning and no warning conditions across the FA, NA, UA, and TA contexts. The top figure represents the muted warning (white) condition and the bottom represents the auditory warning (grey) condition.

While warnings had little effect on younger driver confidence, they shifted ratings for both older and middle-aged drivers. For middle-aged drivers warnings diminished confidence, particularly during the FA and NA contexts. In contrast, warnings increased older driver (both impaired and unimpaired) confidence ratings. As shown in the top of Figure 16, during the no warning condition older drivers had more confidence during FA and NA contexts than during UA and TA contexts, $t(351) \geq 4.70$, $p \leq 0.001$. Noteworthy, when these older driver groups received warnings they became more confident during the UA and TA contexts such that they indicated they were as confident as they were during the FA and NA contexts.

Driver hazard ratings

Hazard ratings represent how hazardous drivers considered the situations they viewed and ranged from not at all hazardous to extremely hazardous. Table 5 provides a summary of the main and interaction effects for hazard ratings. Figure 17 plots mean (±1 S. E.) hazard rating across the different driver groups and different alarm contexts for the warning (grey) and no warning conditions (white). Along with table 7 the figure confirms that drivers distinguish the four alarm contexts. Drivers rated TA (M=1.53, S.E.=0.03) and UA (M=1.36, S.E.=0.04) contexts as more hazardous than FA (M=0.47, S.E.=0.03) and NA contexts (M=0.58, S.E.=0.03).

Table 7. Significant pairwise comparisons for main effects for hazard ratings.

	Estimate (Δ)	t-value	d.f.	95% CI (Adj.)
Group				
Younger vs. Older	0.328	$2.38^{\pm}$	118	(-0.031, 0.687)
Younger vs. Older (UFOV)	0.573	4.05^{**}	118	(0.204, 0.941)
Middle vs. Older	0.325	$2.36^{\pm}$	118	(-0.034, 0.684)
Middle vs. Older (UFOV)	0.569	4.03^{**}	118	(0.201, 0.938)
Alarm Context				
FA vs. NA	-0.115	-3.12^{**}	351	(-0.209, -0.02)
FA vs. UA	-0.891	-24.24^{***}	351	(-0.986, -0.796)
FA vs. TA	-1.048	-32.93^{***}	351	(-1.13, -0.966)
NA vs. UA	-0.776	-21.12^{***}	351	(-0.871, -0.681)
NA vs. TA	-0.933	-29.33^{***}	351	(-0.871, -0.681)
UA vs. TA	-0.157	4.94^{***}	351	(-9.075, 0.239)

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

More importantly, figure 17 demonstrates that hazard ratings for an alarm context depend on who is making the evaluation. Younger, middle-aged, and older drivers (with and without UFOV impairments) considered FA and NA contexts fairly benign regardless of the warning condition. Although the graph suggests that older drivers with UFOV impairments rated these contexts as more hazardous than the other groups, the differences (except in one case) were not significant.

When the UA and TA contexts are considered, four interesting differences arise between the driver groups. First, middle-aged and older drivers rated UA and TA similarly regardless of warning condition. Second, overall older drivers with and without UFOV impairments rated both TA and UA contexts as more hazardous than younger and middle-aged drivers (respectively, older: $t(118) \geq 2.36$, $p \leq 0.1$ and older (UFOV): $t(118) \geq 4.03$, $p \leq 0.01$). However, hazard ratings varied across the groups for these two alarm types. Third, the graph suggests that warnings led younger drivers to increase their hazard ratings for UA and TA contexts. Comparisons indicate that the warning only caused significant shifts in hazard ratings for the TA context. Specifically, when younger drivers received warnings they rated TA contexts as more hazardous than they did without warnings and also as more hazardous compared to UA contexts for both warning and non-warning conditions, $t(351) \geq 3.9$, $p \leq 0.001$. Fourth, The graph suggests a trend for older drivers to consider UA and TA contexts more hazardous when they received warnings. However, unlike the younger drivers the change in hazard ratings did not reach significance after post-hoc comparison adjustments were made.

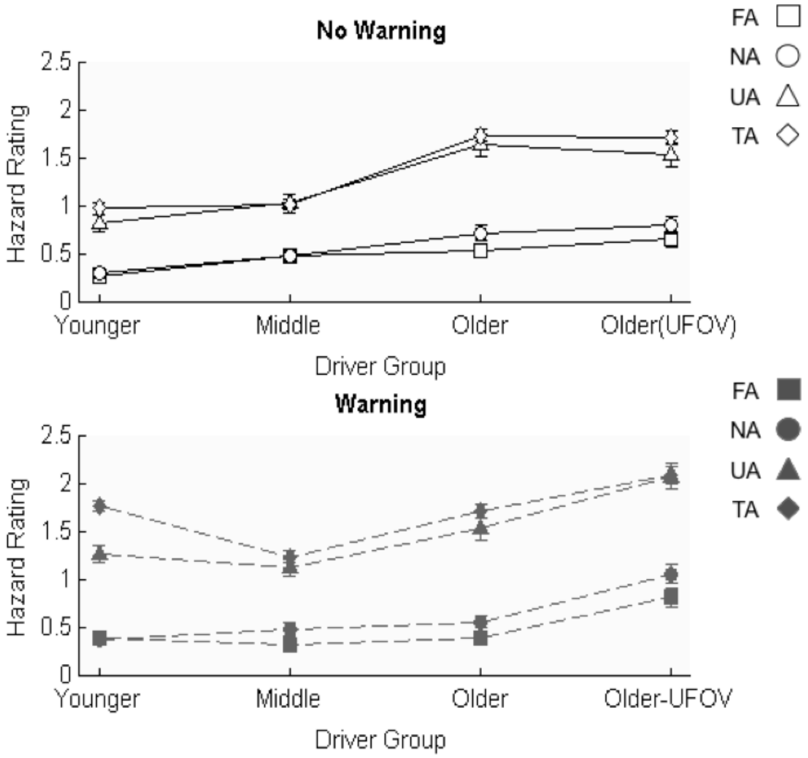

Figure 17. **Mean (±1 S. E.) hazard rating for younger, middle-aged, and older drivers (with and without UFOV impairments) across the FA, NA, UA, and TA contexts. The top figure represents the muted warning (white) condition and the bottom represents the auditory warning (grey) condition.**

Driver responses

Driver responses represent whether or not drivers indicated that they would slow down (by pressing the mouse button) while watching the video. Table 5 provides a summary of the main and interaction effects for driver responses. Figure 18 plots the mean (±1 S. E.) response frequency across the different driver groups and alarm contexts for the warning (grey) and no warning (white) conditions. The figure and table 8 show that as expected drivers responded more during the TA (M=92%, S.E. =

1) and UA (M=92%, S.E.=1) contexts compared to the FA (M=16%, S.E.=1) and NA (M=18%, S.E.=1) contexts.

One of the more interesting differences relates to how the driver groups responded during the UA and TA contexts. While drivers responded 90% or more of the time during the TA contexts, the driver groups did not respond uniformly. Specifically, across the warning and no warning conditions older drivers with UFOV impairments (M = 80%) responded less often during the TA contexts compared to younger (M=98%) and middle-aged drivers (M=98%), t(351)≥3.63, p≤0.001.

Table 8. Significant pairwise comparisons for main effects for response frequency during the different alarm contexts.

	Estimate (Δ)	t-value	d.f.	95% CI (Adj.)
Alarm Context				
FA vs. UA	-0.543	-37.06***	351	(-0.581, -0.505)
FA vs. TA	-0.757	-59.62***	351	(-0.789, -0.724)
NA vs. UA	-0.522	-35.6***	351	(-0.559, -0.484)
NA vs. TA	-0.735	-57.93***	351	(-0.768, -0.702)
UA vs. TA	0.214	16.83***	351	(0.181, 0.246)

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

A somewhat different pattern emerges for the UA contexts. During these situations, younger and older drivers responded less frequently (M=64) compared to middle-aged drivers (M=79, t(351)≥3.46, p≤0.05). The effect of warnings on response frequency during the UA context is particularly interesting. Overall, warnings did not show a strong influence on how drivers responded during the different alarm contexts. An examination of figure 18 reveals that younger driver responded consistently across the warning and no warning conditions during the UA context. However, older drivers with UFOV impairments responded more often during these situations when they were warned. While the difference between the warning and no warning condition was not significant, there is indirect support that the response frequency did increase. Specifically, the difference between the UA context and TA context eroded.

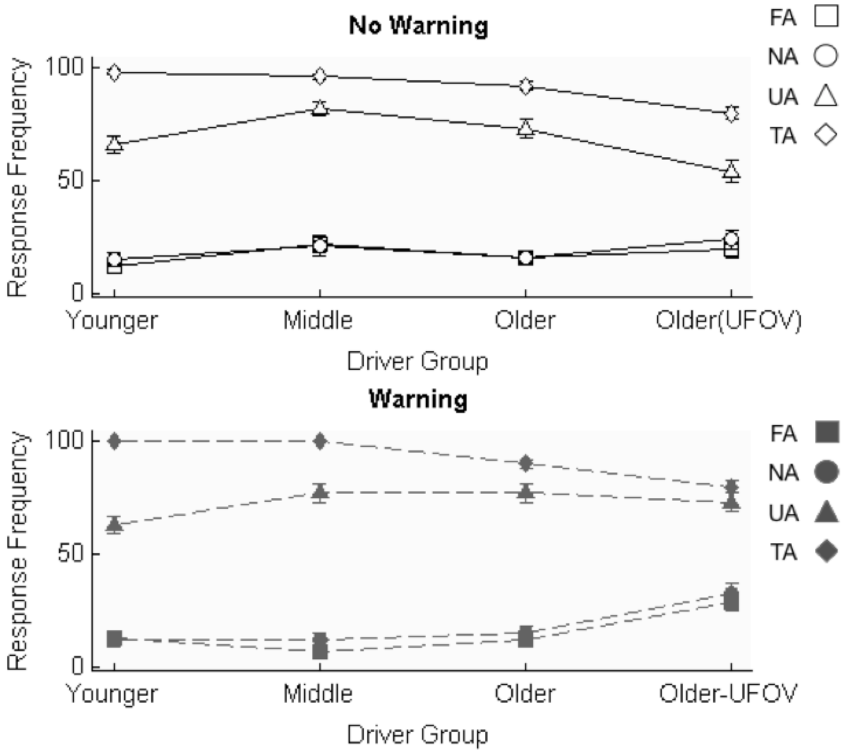

Figure 18. Mean (±1 S. E.) response frequency for younger, middle-aged, and older drivers (with and without UFOV impairments) across the FA, NA, UA, and TA contexts. The top figure represents the muted warning (white) condition and the bottom represents the auditory warning (grey) condition.

Response times to UA and TA contexts were examined to determine when drivers responded to these situations. These times are calculated in relation to the warning algorithm with zero indicative of the driver making a response at the time the system did or would have generated an alarm. Positive reaction times represent situations where the driver responds later than the system. Negative values represent cases where the driver is quicker to respond than the system. Table 9 provides a summary of the statistical main and interaction effects for response times for the UA and TA contexts.

As shown in figure 19 and in table 10, younger, middle-aged and older drivers all responded to UA and TA contexts prior to the system whereas older drivers with UFOV impairments did not, $F(3,118) \geq 6.35$, $p<0.001$. Drivers were also slower to respond during the warning condition compared to the no warning condition ($F(1,118) \geq 3.10$, $p<0.10$).

Table 9. Overview of statistical main and interaction effects for RTs during the UA and TA context.

	Num DF	Den DF	F Value	
			UA Context	TA Context
Group	3	118	6.35^{***}	10.37^{***}
Warning	1	118	5.52^{**}	$0.08^{\pm}$
Group x Warning	3	118	0.76	0.15

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

Table 10. Significant pairwise comparisons for main effects for RTs during the UA and TA context.

	Estimate (Δ)	t-value	DF	95% CI (Adj.)
UA				
Older vs. Older (UFOV)	-0.63	-2.78^{**}	118	(-1.22, -0.04)
Older (UFOV) vs. Younger	0.84	3.67^{**}	118	(0.246, 1.447)
Older (UFOV)vs. Middle-aged	0.91	4.01^{**}	118	(0.317, 1,500)
TA				
Older vs. Older (UFOV)	-0.63	-3.35^{**}	118	(-1.123, -0.141)
Older (UFOV) vs. Younger	0.669	3.57^{**}	118	(0.181 1.159)
Older (UFOV) vs. Middle-aged	1.04	5.53^{***}	118	(0.549, 1.529)

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

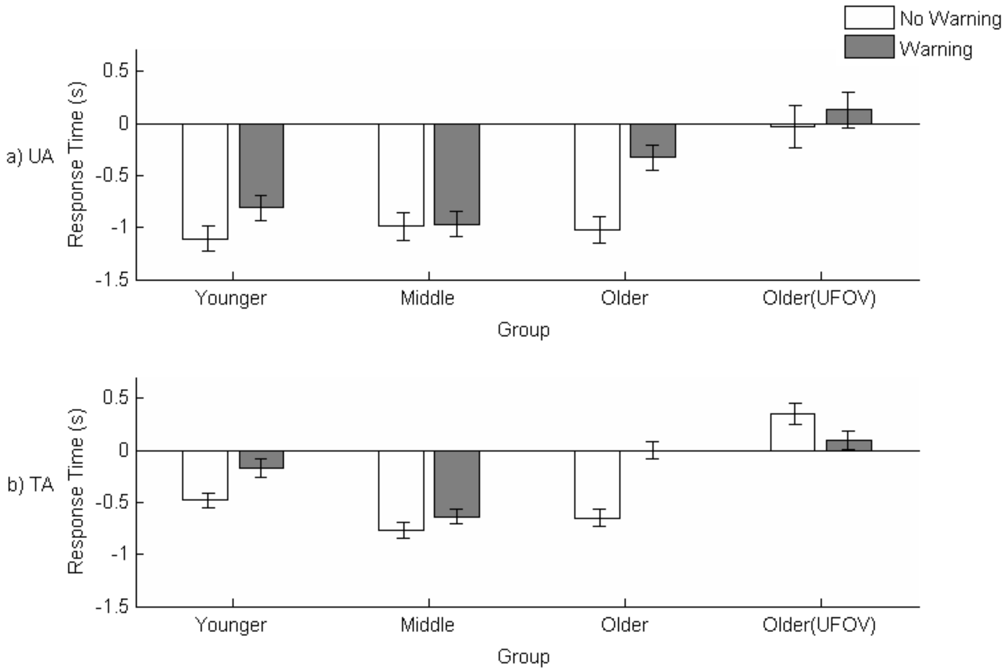

Figure 19. Mean (±1 S. E.) reaction times for UA (figure 19a) and TA (figure 19b) contexts for younger, middle-aged, and older drivers (with and without UFOV impairments) for warning and no warning conditions. Zero corresponds to the time that the system would have or did generate an alarm. Therefore negative values indicate that drivers responded prior to the alarm and positive values indicate that drivers responded after the alarm.

Conclusions

Overall the results confirm that alarm context influences drivers' confidence, hazard evaluation and response patterns. Broadly speaking, drivers consider UA and TA contexts to be more hazardous than FA and NA contexts. However, the evaluation of and pattern of responses during different alarm contexts depends also on the type of driver.

As expected, younger drivers were extremely confident regardless of whether they received warnings or the context of the alarm. Unlike middle-aged and older drivers, warnings did little to modify younger driver confidence. These results suggest that younger driver confidence is extremely resilient. These findings fit nicely with previous research that suggests younger drivers are more confident in their driving abilities compared to other driver groups and express high levels of confidence across different skills and situations (Deery, 1999; DeJoy, 1992; Finn & Bragg, 1986; Matthews & Moran, 1986; McKenna, et al., 1991; Tränkle, et al., 1990).

Various studies have found that responses are conditioned by event severity (Lee, et al., 2002; Lees & Lee, 2007; NHTSA, 2005a). Therefore it is not surprising that younger drivers, like the other driver groups, rated UA and TA contexts as more hazardous than FA and NA contexts. This is consistent with the nature of these different situations and how they were chosen. Younger and middle-aged drivers demonstrated similarities across hazard ratings, responses and response times during the TA contexts. In these situations, both groups responded 99% of the time and tended to respond before the system.

The difference between younger and middle-aged drivers has more to do with how drivers performed during the UA context. While these drivers had similar hazard ratings for these situations younger drivers responded less frequently. Such findings seem to reflect diminished hazard perception ability within younger drivers (Finn & Bragg, 1986; Horswill & McKenna, 2004; McKenna & Crick, 1994; Pelz & Krupat, 1974; Pollatsek, Narayanaan, Pradhan, & Fisher, 2006; Quimby & Watts, 1981). Specifically, younger drivers tend to have lower risk estimates and to be less accurate and slower in perceiving hazards (Horswill & McKenna, 2004; McKenna & Crick, 1994). While good performance during the TA contexts may seem at odds with diminished hazard perception it is not. Specifically, typical hazard perception tasks are designed in a way that drivers must anticipate potential hazards not infer them directly. As such, the UA contexts are more consistent with the types of events usually included in hazard perception tasks. In contrast, TA events are more representative of simple RT tasks where drivers respond to a simple cue in the environment (e.g., brake lights). Taken together, these findings suggest younger drivers are extremely confident and may fail to identify potentially hazardous situations even when they receive warnings. As such, warning systems may fail to modify how these drivers evaluate and respond to hazards within their environment.

Older drivers with and without UFOV impairments were less confident than the younger drivers which is consistent with other research (G. Ho, et al., 2005). However, confidence in older drivers was influenced not only by the alarm context but also by system warnings. Specifically, warnings increased confidence for older drivers to the extent that these drivers became more confident during the most demanding situations (UA and TA context). Unfortunately, for older drivers with UFOV impairments this increased confidence did not coincide with improved performance when drivers were presented with the UA and TA contexts.

Drivers with UFOV impairments considered UA and TA contexts to be more hazardous than younger and middle-aged drivers yet they responded less frequently and slower during both situations. These drivers took much longer to make a response and were the only group that the system outperformed. Given that the hazard ratings of these drivers matched those of older drivers another possibility is that these drivers were unable to generate a response within the small time window. These findings are consistent with other studies that have found diminished hazard perception ability in older drivers (Horswill, et al., 2008; Horswill, et al., 2009). Horswill et al. (2008) found that diminished contrast sensitivity, greater UFOV impairments and slower simple RT performance was associated with poorer hazard perception ability. Slower responses seem to have been particularly problematic and reflect the tendency of older drivers to require more time to process information, make a decision, and implement a response (Lerner, 1994).

Noteworthy, older drivers with UFOV impairments benefited from warnings during the UA context in that they responded more often when warned. Such benefits may explain why confidence ratings increased during the UA context when these drivers received warnings. These results suggest that older drivers with UFOV impairments may benefit from warnings by aiding bottom up processing and response execution.

In most other cases the drivers demonstrated similar responses across the warning and no warning condition. A number of explanations might account for why warnings did little to increase compliance or reduce response times during the UA and TA contexts. First, drivers received a large number of alarms across the experiment. Therefore, the alarms might have lost their wow factor. Second, the overall reliability for three of the four systems was extremely low. The failure to benefit from warnings might have led drivers to rely more on their own judgement when determining how and when to respond (Getty, et al., 1995; Parasuraman & Riley, 1997; Wickens & Dixon, 2007). Third, literature suggests that the benefit of such systems relates most to the ability to direct attention. Therefore, greater benefits might have been observed if drivers were required to divide their attention between driving and another task. Fourth and most likely is that drivers respond to the situation not the warning. Other driving studies have found that drivers adopt similar response patterns across warning and no-warning conditions and responses are conditioned by event severity (Lee, et al., 2002; Lees & Lee, 2007; NHTSA, 2005a).

CHAPTER 5.

ALARM CONTEXT INFLUENCES TRUST

This chapter examines how younger, middle-aged, and older drivers perceive alarms that occur within different alarm contexts. Trust data was analyzed using a mixed model procedure in SAS (Statistical Analysis System) version 9.2. For the analysis, group was included as a between-subjects variable and alarm (FA, NA, UA, and TA) was treated as a within-subjects variable repeated across each block. Subject was represented as a random effect and the compound symmetry covariance structure was used. Pairwise comparisons were computed with Tukey-Kramer adjustments for outcomes with significance level of 0.05 or better. A profile analysis was conducted using the repeated measures module under the general linear model in SPSS v17. The analysis required that alarm perception ratings be treated as the repeated measure. Therefore, data across similar trials was collapsed and a separate analysis were conducted for each of the alarm (FA, NA, UA, and TA) with group as a between-subjects measure.

Trust in different alarm types across trials

As shown by Figure 22, trust ratings substantially differed across failure and non-failure trials. Overall, drivers are fairly consistent in their ratings of similar trial types. For example, across the different blocks drivers rated TA contexts as highly trustworthy. The figure demonstrates that drivers differentiated between FAs, NAs, UAs, and TAs. Similar to other studies examining unreliable automation, the graph also suggests that driver trust was able to recover from errors when the system provided accurate information (Abe & Richardson, 2006a; Breznitz, 1983; Kantowitz, et al., 1997; Lee & Moray, 1992).

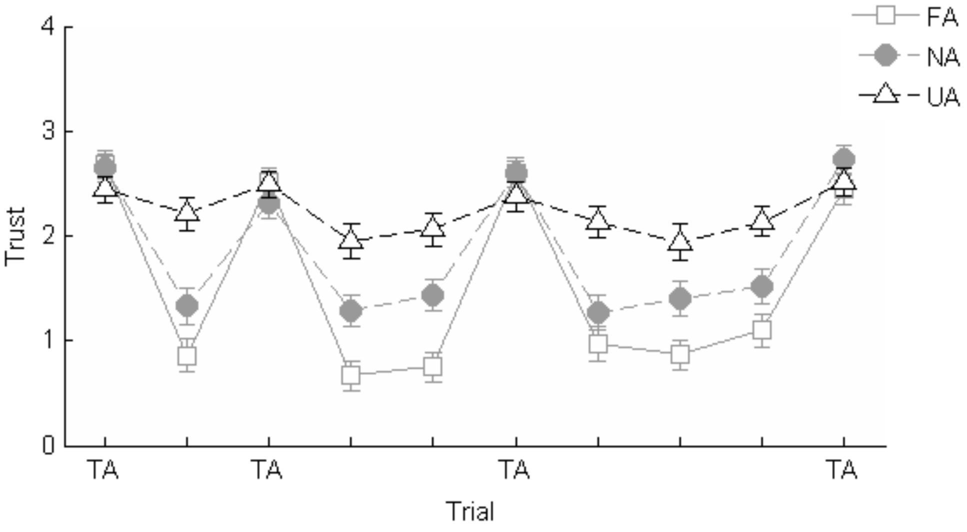

Figure 20. Trust ratings during the FA, NA and UA blocks. Trials with TA on the axis indicate a TA trial. All other trials represented one of the three types of alarms.

Table 11 provides an overview of the main and interaction effects for trust ratings. Figure 21 plots the mean (±1 S. E.) trust rating for each driver group according to the alarm context. At a high level, trust appears to depend on the context of the alarm with drivers trusting TAs (M: 2.52, S. E.0.03) more than UAs (M: 2.11, S. E.0.05), NAs (M: 1.35, S. E.0.06) and FAs (M: 0.08, S. E.0.05). However, as shown trust depends on the combination of driver and alarm, *F*(3,351)=8.06, p<0.001.

Unlike the other alarms, all drivers considered TAs to be relatively trustworthy. However, older drivers with UFOV displayed a different pattern of trust across the other alarm types compared to younger and older drivers. As shown by figure 21, trust ratings for older drivers with UFOV impairments were less sensitive to alarm context. Specifically, the range in their trust ratings was 1.66 compared to 1.83 for younger drivers, 1.88 for older drivers, and 1.88 for middle-aged drivers.

Overall, the results suggest that alarm context matters and drivers consider some alarms to be more trustworthy than others. Specifically, most drivers trust TAs and UAs more than FAs and NAs. A notable exception was older drivers with UFOV impairments who did not differentiate TAs and UAs from NAs. It is also interesting that unlike middle-aged and older drivers with impairments, younger drivers did not differentiate FAs and NAs.

Table 11. Overview of statistical main and interaction effects for trust ratings.

Effect	Num DF	Den DF	F- Value
			Trust
Group	3	60	$2.39^{\pm}$
Alarm Context	3	177	390.41^{***}
Group x Alarm Context	9	177	8.06^{***}

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

Table 12. Significant pairwise comparisons for main effects for trust ratings.

	Estimate (Δ)	t-value	DF	95% CI (Adj.)
Alarm Context				
FA vs. NA	-0.47	-7.82^{***}	177	(-0.63, -0.32)
FA vs. UA	-1.23	-20.29^{***}	177	(-1.38, -1.10)
FA vs. TA	-1.63	-31.26^{***}	177	(-1.77, -1.50)
NA vs. UA	-0.75	-12.47^{***}	177	(-0.91, -0.60)
NA vs. TA	-1.16	-22.24^{***}	177	(-1.30, -1.03)
TA vs. UA	0.41	7.84^{***}	177	(0.27, 0.54)

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

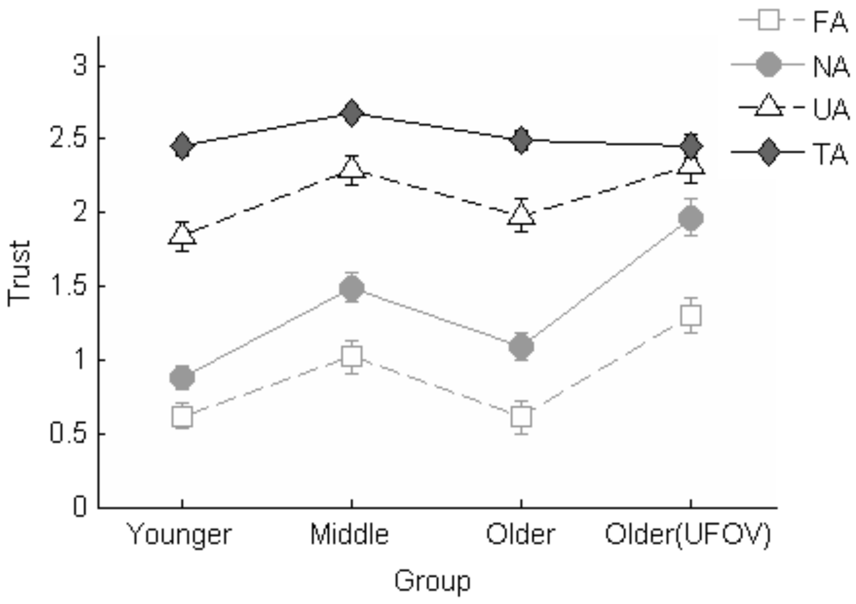

Figure 21. **Mean (±1 S. E.) trust rating for younger, middle-aged, and older drivers (with and wihtout UFOV imapairments) for FAs, NAs, UAs, and TAs.**

Trust, understandability, usefulness, and appropriateness ratings

A similar set of findings emerged for the alarm perception data. Therefore, a profile analysis was carried out to compare trust, understandability, usefulness, and appropriateness ratings between the younger, middle-aged, and older (with and without UFOV impairments) drivers for the different alarm types. The goal of the analysis was to determine the following: 1) are the groups parallel between the different measures (table 13)?, 2) are the groups at equal levels across the different measures (table 14)?, 3) do the profiles exhibit flatness across the different measures (table 13)?

Table 13. Overview of statistical between-subjects effects for ratings during the FA, NA, UA, and TA context.

Effect	Num DF	Den DF	F- Value			
			FA	NA	UA	TA
Intercept	1	59	121.85^{***}	146.38^{***}	439.58^{***}	982.84^{***}
Group	3	59	6.49^{***}	5.60^{***}	$2.64^{\pm}$	$0.72^{\pm}$

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

Table 14. Overview of statistical within-subjects effects for ratings during the FA, NA, UA, and TA context.

Effect	Num DF	Den DF	F- Value			
			FA	NA	UA	TA
Subjective Measure	3	177	11.23^{***}	21.51^{***}	29.84^{***}	29.27^{***}
Subjective Measure x Group	9	177	0.98	0.67	1.23	0.30

±p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001

Figure 22 presents the profile plots for the FAs, NAs, UAs, and TAs. The plots show parallelism between the different measures across the different groups, F(9, 59) ≤0.131, ns). Differences between the profiles suggest that drivers rated FAs and NAs differently than UAs and TAs. For all alarms the profiles did not exhibit a flatness across the different measures, F(3, 59)≥11.23, p≤0.001. Inspecting the profile plots suggests that drivers indicated similar levels for trust and understandability and similar ratings for usefulness and appropriateness for FAs and NAs. In contrast drivers indicated similar levels for trust and appropriateness for UAs and TAs. Drivers had the highest ratings for understandability during these situations.

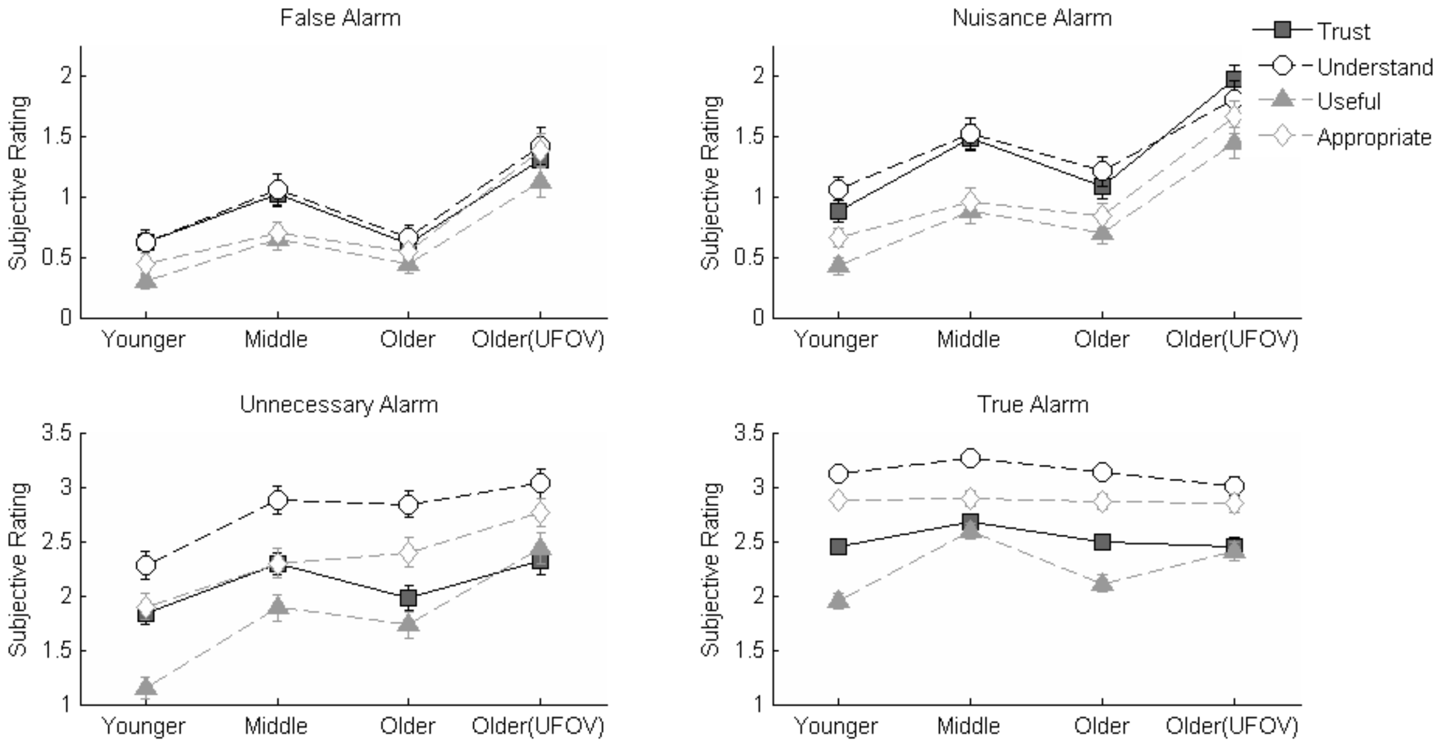

Figure 22. Mean (±1 S. E.) alarm perception ratings rating for younger, middle-aged, and older drivers (with and without UFOV impairments for FA, NA, UA, and TA contexts.

One of the most interesting implications of the analysis is how the driver groups differ in their ratings across the different alarms. Significant differences between the groups were found for the FAs, NAs, and UAs but not for the TAs. Older drivers with UFOV impairments rated FAs and NAs more positively compared to the other driver groups $F(3, 59) \geq 4.59$, $p \leq 0.01$. There was also a marginally significant difference between younger and older drivers with UFOV impairments for UAs. It seems that younger drivers considered these warnings less useful and understandable than the other driver groups. These findings suggest that older drivers with impairments may be more forgiving of system failures.

Conclusions

The primary goal of this research was to determine how different driver groups evaluate collision warning system alarms that occur in different driving contexts. Consistent with the findings of Lees and Lee (2007) driver trust depends on alarm context. When driver trust is considered a hierarchy of alarms is revealed with unnecessary alarms being least detrimental to driver trust and false alarms being most detrimental. However, trust depends not only on the alarm context but also on the driver making the evaluation.

The results from Chapter 4 suggest that drivers respond uniformly to FA and NA contexts. However, alarm perception data reveals that drivers trust FAs more than NAs. While several studies have found that older drivers may be more trusting of unreliable automation compared to younger drivers (Fox & Boehm-Davis, 1998; G. Ho, et al., 2005; Lee, Gore, & Campbell, 1999; NHTSA, 2005a) the current study found this to hold only for older drivers with UFOV impairments. Specifically, older drivers with UFOV impairments were less negative when rating FAs and NAs.

Interestingly, these drivers had less variability in their ratings of the different alarm types. This suggests that trust calibration differs across driver groups.

While similar findings occurred for the various alarm perception rating data, the results do indicate that these measures differ from each other. Drivers indicated similar levels for trust and understandability and similar ratings for usefulness and appropriateness for FAs and NAs. In contrast, drivers indicated similar levels for trust and appropriateness for UAs and TAs. Drivers had the highest ratings for understandability during these situations. The lower understandability and usefulness ratings for UAs made by younger drivers seems to confirm diminished hazard perception ability during these situations (Finn & Bragg, 1986; Horswill & McKenna, 2004; McKenna & Crick, 1994; Pelz & Krupat, 1974; Pollatsek, Narayanaan, et al., 2006; Quimby & Watts, 1981). Unlike the other alarm types, the different driver groups converged in their positive ratings of TAs for trust, understandability, usefulness and appropriateness. Unlike the UAs all drivers seemed to indicate a high degree of understandability for why these alarms occurred.

Overall, these findings provide support that not all alarms have the same influence and that researcher’s should look beyond signal detection theory (Allendoerfer, et al., 2008; Barnes, et al., 2007; Friedman-Berg & Allendoerfer, 2008; Lees & Lee, 2007; Seagull & Sanderson, 2001; Woods, 1995). Specifically, the alarm context and the evaluator influence how tolerant people are of warnings. Similar to Smith and Kallhammer (2010) this study found subjective criteria might be particularly useful when evaluating and designing such systems. (Smith & Kallhammer, 2010)

CHAPTER 6.

GENERAL CONCLUSIONS

Taken together this study confirms that alarm context matters and that not all alarms are detrimental to trust and compliance. However, the evaluation and response to different alarm contexts depends upon the driver making the evaluation. Specifically, there are differences in how at-risk populations (younger and older drivers with UFOV impairments) evaluate and respond to certain alarm contexts.

Alarm context matters

Similar to the Automotive Collision Avoidance Field Operational Test (ACAS FOT) results (NHTSA, 2005a), drivers in this study adapted their response to the alarm context. In both studies drivers responded less frequently to out-of-host path conflict scenarios (false alarm - FA and nuisance alarm -NA contexts) than they did to unnecessary alarm (UA) or true alarm (TA) contexts that involve a transitioning or in-host conflict scenarios. While the response data suggests that FA are similar to NA and UA are similar to TA, greater demarcation exists among the different alarm contexts when subjective data is considered. Specifically, when trust, hazard ratings and alarm perception data are considered the following hierarchy emerged (from worst to best) FA, NA, UA, then TA.

While drivers adopted similar response patterns during the FA and NA contexts, differences exist in how drivers perceived these out-of-host path conflict scenarios. Drivers considered NA contexts to be more hazardous than FA contexts. Additionally, middle-aged and older drivers trusted NAs more than FAs suggesting that for most drivers FAs are more detrimental than NAs. As previously shown UA and TA were less detrimental (Lees & Lee, 2007).

Drivers differentiated between UA and TA contexts. UA contexts involving a transitioning host path conflict are considered to be less hazardous than TA contexts involving in-host path conflict scenarios. As such drivers respond more frequently during TA contexts and are more trusting of alarms that occur within these contexts. These findings support the general framework provided by Lees and Lee (2007) in differentiating alarm types. They are also consistent with studies showing the drivers modulate their response according to event severity and conflict type (Lee, et al., 2002; NHTSA, 2005a). However, as will be discussed in the next section the evaluation of and pattern of response during these alarm contexts depend on who the driver is.

Younger drivers respond less frequently during the UA contexts

Younger drivers were extremely confident during the different alarm contexts and unlike some of the other groups were unaffected by warnings. While younger and middle-aged drivers evaluated and responded to TA contexts in a similar manner the same was not true of performance during UA contexts. Specifically, despite rating UA contexts as hazardous as middle-aged drivers, younger drivers responded less frequently during these situations. These results are consistent with other research that has shown that younger drivers overestimate their driving skill and perceive lower levels of risk compared to middle-aged experienced drivers (Deery, 1999; DeJoy, 1992; Finn & Bragg, 1986; McKenna, et al., 1991; Tränkle, et al., 1990). The findings of the current study are also consistent with studies showing that younger drivers often fail to appreciate or recognize risk or hazardous situations (Finn & Bragg, 1986; Horswill & McKenna, 2004; McKenna & Crick, 1994; Pelz & Krupat, 1974; Pollatsek, Narayanaan, et al., 2006; Quimby & Watts, 1981). Diminished responses during UA context likely reflect diminished hazard perception ability where these drivers unlike their middle-aged counterparts failed to recognize or anticipate the potential hazards within these situations.

Older drivers with UFOV impairments respond less frequently and more slowly than other driver groups during the UA and TA contexts

Older drivers with UFOV impairments differed in how they viewed FAs and NAs. Several studies suggest that older drivers may be more tolerating of system failures compared to other age groups (Dingus, et al., 1997; Fox & Boehm-Davis, 1998; NHTSA, 2005a). For example, in the ACAS FOT older drivers rated the collision warning system more favourably compared to younger and middle-aged drivers who complained of FA and NA. Fox and Boehm-Davis (1998) found that older drivers trusted unreliable advanced traveller information systems more than younger drivers.

In the current study, younger, middle-aged and older drivers indicated similar levels of trust for FAs and NAs. In contrast, older drivers with UFOV impairments were more positive towards these alarms in terms of trust, understandability, usefulness and appropriateness. This is somewhat surprising given that older drivers with and without impairments were not significantly different in terms of confidence, hazard ratings or response rates. The difference might reflect an ability of trust to be more sensitive than these other measures.

While older drivers considered UA and TA contexts to be more hazardous than younger and middle-aged drivers they often failed to elicit an appropriate response during these situations compared to middle-aged and older drivers. In part, such differences might reflect diminished hazard perception ability within these populations. For example, several studies have shown that drivers with UFOV impairments are less accurate and slower in perceiving hazards (Horswill, et al., 2008;

Horswill, et al., 2009). Unlike the other groups older drivers with UFOV impairments were unable to identify and execute a response prior to the systems warnings during both the UA and TA contexts. Such diminished performance likely reflects declines in cognition, diminished perception, and slower responses (Horswill, et al., 2008; Horswill, et al., 2009; Lerner, 1993, 1994). The inability of these drivers to respond within the short time frame seems particularly relevant given that if drivers fail to plan and execute responses in a timely manner they will crash.

The role of warnings in collision avoidance

The goal of collision warnings is to mitigate crash involvement by directing the driver's attention to hazards. However, in this study warnings often failed to influence whether or when drivers responded. Unlike previous research on such systems (Dingus, et al., 1997; C. Ho, et al., 2007; Kramer, et al., 2007; Lee, et al., 2002; Scott & Gray, 2008), there was not RT benefit in the current study. Overall this probably reflects that most drivers picked up cues prior to receiving a warning. However, it seems that when warnings were provided drivers became more reactive than proactive in seeking out information. While drivers did not always comply with warnings, they relied on the system to determine when to respond during the UA and TA contexts (Meyer, 2004). Reliance refers to how operators respond when the system indicates no hazard, whereas compliance describes the operators response when the system detects a hazard and issues an alert (Dixon & Wickens, 2006; Meyer, 2004). The RT data suggest that younger and older drivers were slower to respond when they received warnings. Overall, older drivers with UFOV impairments took longer to respond during both the warning and no warning condition.

Overall, warnings also had little effect on response frequency. In some respects these results are not surprising considering research suggesting that drivers respond to the situation, not the alarm. That is, drivers do not respond if they do not agree with the warning. For example, in the ACAS FOT warnings did not change how drivers responded (NHTSA, 2005a). Warnings may have failed to benefit drivers because their attention was already on the road and therefore they were already able to identify and respond to roadway hazards regardless of whether they received an alarm. Another reason that warnings may have failed to benefit these drivers is due to the timing of the alarm.

Despite warnings not influencing most drivers, older drivers with UFOV impairments that received warnings responded more frequently to the UA situation. This suggests that warnings might benefit these drivers by helping them identify situations where a response might be needed. At the same time caution is needed in applying these results since there were no differences in response times. The warnings also failed to benefit these drivers during the TA context. Earlier warnings might have induced a reaction time benefit for drivers with attention impairments who were slow to respond to TA contexts. Given such a limited effect on responses it is somewhat worrisome that warnings increased confidence for these drivers during the TA and UA contexts.

Summary of key findings

1. Alarm context matters. Response data clearly differentiated FA and NA contexts from UA and TA contexts (FA = NA, UA = TA). Greater differentiation exists when subjective data was considered and suggests that drivers clearly distinguish between the different alarm contexts (FA<NA<UA<TA).
2. Drivers evaluated and responded similarly across warning and no warning conditions.

3. Younger drivers indicated a high degree of confidence across the different conditions. Diminished hazard perception ability and lower risk perception likely account for the fewer responses these drivers made during the UA context.
4. Older drivers were less confident than younger drivers. While older drivers with and without attention impairments indicated similar hazard ratings for UA and TA contexts, older drivers with UFOV impairments responded less frequently during these situations. Diminished hazard perception ability, slower simple response times, degraded contrast sensitivity likely account for the fewer and slower responses. Unlike the other groups, these drivers benefit from warnings for the UA context.

Implications for future research

While some might argue that only responses matter, subjective opinions have a strong influence on how these systems influence drivers. Similar to recent research this study demonstrates that there is much utility in using subjective criteria to evaluate and inform design so that these types of systems match the drivers experience (Smith & Kallhammer, 2010). Relating the findings back to signal detection framework it would seems that FA and NA are most characteristic of a FA, whereas UA and NA are most characteristic of a hit. However, such a framework still fails to capture differences between these alarms. Unlike response data, subjective data provide a window into whether drivers can differentiate alarms and what alarms might be most detrimental to drivers over time.

Overall the findings suggest that drivers can and do differentiate alarm contexts. Specifically, FAs are more detrimental than NAs which are more detrimental than UAs. This suggests that cases where the operator is unable to identify the cause or source of alarm (FA) will be most destructive to trust. A greater degree of trust will result when the operator is able to identify the cause or source of the alarm because the alarms have a consistent causal factor (NA). Finally, alarms that occur in situations where the operator can identify the source and that relate to potential hazards are most resilient and in some cases might benefit the operator (UA). The results also suggest that operators with different levels of experience or different characteristics may differ in how they evaluate alarms. Such difference may erode for TA contexts. For example in the current study, the different driver groups made similar evaluations for the TA context but differed in their ratings of FA, NA, and UA.

This research provides additional support for the framework developed by Lees and Lee (2007). Such classifications allow a deeper understanding of how different alarms contribute to dissatisfaction. Researchers are encouraged to look beyond signal detection theory when trying to capture the relationship between human automation interaction (Allendoerfer, et al., 2008; Barnes, et al., 2007; Friedman-Berg & Allendoerfer, 2008; Lees & Lee, 2007; Seagull & Sanderson, 2001; Woods, 1995). At the same time greater benefits might be obtained if researchers used multiple perspectives. For example, engineering approaches that identify conflict types (e.g. the conflicts identified in the ACAS FOT study, NTHSA, 2008) might be further differentiated using more of a user perspective (e.g. the differentiation of alarms by Lees and Lee, 2007) to identify areas of exploration that relate to how real-world systems can and do operate. Such research should try to quantify subjective ratings for various scenarios to develop cut-offs between acceptable and unacceptable alarms. Such an approach was recently taken when researchers evaluated how post encroachment time (PET) influence drivers acceptance of left-turn encroachment alerts (Smith & Kallhammer, 2010). The study found that drivers were more accepting of alarms where the PET interval was less than 2.2 seconds.

Limitations and considerations for future research

Currently, researchers use simulators or instrumented vehicles to examine collision warning systems. The hazard perception task offers an alternative medium to evaluate collision warning system design. There are several benefits to this approach but the main benefit is that it allows drivers to be evaluated under the exact same conditions. This approach can allow designers to rule out design alternatives because researchers are using the same testing conditions and therefore differences observed relate to manipulations being made. This approach would allow designers to use the data they collect during testing of algorithms. However, a number of improvements could be made to increase task validity and realism.

The software used did not allow drivers to respond using a brake pedal. Using such an input device is preferable because it is congruent with how drivers respond in the real world and would minimize the need for computer familiarity. In addition, using a brake pedal might incorporate additional context to reinforce to the participant that they following the situation being presented within the videos. This may be particularly important when evaluating older drivers. On the other hand, when examining complex scenes the use of a touch screen or mouse may allow researchers to better understand what drivers consider hazardous about the situation. A limitation of having people indicate the hazard location is that it may delay response times, especially for older drivers.

The current task used low resolution videos collected as part of a field operational task. A major benefit to using these videos is that the timing of the alarm corresponded to an actual system. In the future, it is recommended that such tasks incorporate colour videos collected at a higher frame rate over a longer period of time. This is especially true when using these types of tasks to evaluate populations that might take longer to respond or who have degradations in their vision.

The timing of the warnings might be modified to incorporate a buffer. In the current task, drivers may have responded due to the onset of the lead vehicle brake lights. It is likely that in the real world or in simulated tasks drivers may monitor the situation prior to implementing a response. However, in the current task where the time window was quite short and the preview time was very limited drivers seem to have relied heavily on brake cues. As a result the onset of the warning may appear late. Using a buffer to incorporate these differences, or increasing the preview time might make these tasks better equipped to evaluate warnings.

The current experiment did not incorporate a 100% accurate system where only TAs were used. Such a system would represent a better comparison to evaluate other conditions but was not possible because only a few dozen TA videos were available.

While this methodology is incredibly useful it is also important to note that different tasks might answer different research questions. Using a multitude of platforms researchers may be better able to understand what factors influence drivers when interacting with these systems (Lees, Cosman, Lee, Rizzo, & Fricke, In press). For example, cognitive science paradigms might be most useful in identifying warning signal parameters are most effective at cueing attention (Lees, et al., 2009). At the same time, tasks such as hazard perception tasks and simulated experiments allow researchers to evaluate understand how drivers respond to warnings in more complex environments. Finally, simulators, instrumented vehicles, and naturalistic studies allow researchers to understand how drivers impact and are impacted by their surroundings. Using a variety of methods is likely to be most beneficial to researchers gaining a more complete understanding of how such systems influence drivers.

REFERENCES

Abe, G., & Richardson, J. (2004). The effect of alarm timing on driver behaviour: an investigation of differences in driver trust and response to alarms according to alarm timing. *Transportation Research Part F, 7*, 307-322.

Abe, G., & Richardson, J. (2005). The influence of alarm timing on braking response and driver trust in low speed driving. *Safety Science, 43*(9), 639-654.

Abe, G., & Richardson, J. (2006a). Alarm timing, trust and driver expectation for forward collision warning systems. *Applied Ergonomics, 37*(5), 577-586.

Abe, G., & Richardson, J. (2006b). The influence of alarm timing on driver response to collision warning systems following system failure *Behaviour and information technology, 25*(5), 443-452.

Allendoerfer, K. R., Pai, S., & Friedman-Berg, F. J. (2008). *The complexity of signal detection in air traffic control alert situations*. Paper presented at the Human Factors and Ergonomics Society 52nd Annual Meeting.

Ankrum, D. (1992). Smart vehicles, smart roads. *Traffic Safety, 92*(3), 6-9.

Balakrishnan, J. D., MacDonald, J. A., Busemeyer, J. R., & Lin, A. (2007). The next logical step in the evolution of signal detection analysis. *http://www.cogs.indiana.edu/publications/techreps2002/248/248.pdf*.

Ball, K., & Owsley, C. (1993). The useful field of view test: A new technique for evaluating age-related declines in visual function. *Journal of the American Optometric Association, 64*, 71-79.

Ball, K., Owsley, C., Stalvey, B., Roenker, D. L., Sloane, M. E., & Graves, M. (1998). Driving avoidance and functional impairment in older drivers. *Accident Analysis and Prevention, 30*(3), 313-322.

Barnes, L. R., Gruntfest, E. C., Hayden, M. H., Schultz, D. M., & Benight, C. (2007). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting, 22*, 1140-1147.

Ben-Yaacov, A., Maltz, M., & Shinar, D. (2002). Effects of an in-vehicle collision avoidance warning system on short- and long-term driving performance. *Human Factors, 44*(2), 335-342.

Bisantz, A. M., & Pritchett, A. R. (2003). Measuring the fit between human judgments and automated alerting algorithms: A study of collision detection. . *Human Factors, 45*(2), 266-280.

Bliss, J. P., & Acton, S. A. (2003). Alarm mistrust in automobiles: how collision alarm reliability affects driving. *Applied Ergonomics, 34*(6), 499-509.

Bliss, J. P., Dunn, M., & Fuller, B. S. (1995). Reversal of the cry-wolf effect: An investigation of two methods to increase alarm response rates. *Perceptual and Motor Skills, 80*, 1231-1242.

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behavior in response to alarms of varying reliability. *Ergonomics, 38*(11), 2300-2312.

Breznitz, S. (1983). *Cry-wolf: The psychology of false alarms*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, T. L., Lee, J. D., & McGehee, D. V. (2001). Human performance models and rear-end collision avoidance algorithms. *Human Factors, 43*(3), 462-482.

Bustamante, E. A., Bliss, J. P., & Anderson, B. L. (2007). Effects of varying the threshold of alarm systems and workload on human performance. *Erognomics, 50*(7), 1127-1147.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention, 40*(4), 1282-1293.

Charlton, J. L., Oxley, J., Fildes, B., & Les, M. (2001). *Self-regulatory behaviour of older drivers*. Paper presented at the Road Safety Research, Policing and Education Conference.

Chugh, J. S., & Caird, J. K. (1999). *In-vehicle train warnings (ITW): The effect of reliability and failure type on driver perception response time and trust*. Paper presented at the Human Factors and Ergonomics Society 43rd Annual Meeting, Houston, Texas, USA.

Cosman, J. D., Lees, M. N., Vecera, S. P., Lee, J. D., & Rizzo, M. (In Revision). Inefficient attentional shifting drives age-related impairments in the useful field of view.

de Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human-Computer Studies, 58*(6), 719-735.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research, 30*(4), 225-236.

DeJoy, D. M. (1992). An examination of gender differences in traffic accident risk perception. *Accident Analysis & Prevention, 24*(3), 237-246.

Dingus, T. A., McGehee, D. V., Manakkal, N., Jahns, S. K., Carney, C., & Hankey, J. M. (1997). Human factors field evaluation of automotive headway maintenance/collision warning devices [Special Section]. *Human Factors, 39*(2), 216-229.

Dixon, S. R., & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors, 48*(3), 474-486.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors, 49*(4), 564-572.

Donmez, B., Boyle, L. N., & Lee, J. D. (2006). Drivers' attitudes toward imperfect distraction mitigation strategies. *Transportation Research Part F: Traffic Psychology, 9*(6), 387-398.

Edwards, J. D., Vance, D. E., Wadley, V. G., Cissell, G. M., Roenker, D. L., & Ball, K. (2005). Reliability and validity of useful field of view test scores as administered by a personal computer. *Journal of Clinical and Experimental Neuropsychology, 27*, 529-543.

Evans, L., & Wasielewski, P. (1982). Do accident-involved drivers exhibit riskier everyday driving behavior? *Accident Analysis & Prevention, 14*(1), 57-64.

Evans, L., & Wasielewski, P. (1983). Risky driving related to driver and vehicle characteristics. *Accident Analysis & Prevention, 15*(2), 121-136.

Farber, E., & Paley, M. (1993). Using freeway traffic data to estimate the effectiveness of rear end collision countermeasures *Third Annual IVHS America Meeting*. Washington, DC: IVHS America.

Finn, P., & Bragg, B. W. E. (1986). Perception of the risk of an accident by young and older drivers. *Accident Analysis & Prevention, 18*(4), 289-298.

Fisher, D. L., Pollatsek, A. P., & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention, 12*, 25-29.

Fox, J. M., & Boehm-Davis, D. A. (1998). Effects of age and congestion information accuracy of advanced traveler information on user trust and compliance. *Transportation Research Record, 1621*, 43-49.

Friedman-Berg, F. J., & Allendoerfer, K. R. (2008). *Nuisance alerts in operational ATC environments: Classification and frequency*. Paper presented at the Human Factors and Ergonomics Society 52nd Annual Meeting, New York, NY.

Gao, J., & Lee, J. D. (2006). Extending the decision field theory to model operators' reliance on automation in supervisory control situations. *IEEE Systems, Man, and Cybernetics, 36*(5), 943-959.

Getty, D. J., Swets, J. A., Prickett, R. M., & Gonthier, D. (1995). System operator response to warnings of danger: A laboratory investigation of the effects of the predictive value of a warning on human response time. *Journal of Experimental Psychology: Applied, 1*(1), 19-33.

Goode, K. T., Ball, K., Sloane, M., Roenker, D. L., Roth, D. L., Myers, R. S., et al. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings, 5*(4), 425-438.

Graham, R. (1999). Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics, 42*(9), 1233-1248.

Grayson, G. B., & Sexton, B. F. (2002). *The development of hazard perception training* (No. TRL558): TRL.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Groeger, J. A., & Brown, I. D. (1989). Assessing ones own and others driving ability: Influences of sex, age, and experience *Accident Analysis and Prevention, 21*(2), 155-168.

Grounds, C. B., & Ensing, A. R. (2000). Automation Distrust Impacts on Command and Control Decision Time. *Proceedings of the IEA2000/HFES 2000 Congress*, 3-855 - 853-858.

Gupta, N., Bisantz, A. M., & Singh, T. (2002). The effects of adverse condition warning system characteristics on driver performance: an investigation of alarm signal type and threshold level. *Behaviour & Information Technology, 21*(4), 235-248.

Hagenouw, R. (2007). Should we be alarmed by our alarms? *Current Opinion in Anaesthesiology, 20*, 590-594.

Ho, C., Reed, N., & Spence, C. (2006). Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis & Prevention, 38*, 988-996.

Ho, C., Reed, N., & Spence, C. (2007). Multi-sensory in-car warning signals for collision avoidance. *Human Factors, 49*, 1107-1114.

Ho, G., Wheatley, D., & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers, 17*(6), 690-710.

Horswill, M. S., Marrington, S. A., McCullough, C. M., Wood, J., Pachana, N. A., McWilliam, J., et al. (2008). The hazard perception ability of older drivers. *Journal of Gerontology: Psychological Sciences, 63*, 212-218.

Horswill, M. S., & McKenna, F. P. (1999). The development, validation, and application of a video-based technique for measuring an everyday risk-taking behavior: Drivers' speed choice. *Journal of Applied Psychology, 84*(6), 977-985.

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A*

Cognitive Approach to Situation Awareness (pp. 155-175). Aldershot, UK: Ashgate.

Horswill, M. S., Pachana, N. A., Wood, J., Marrington, S. A., McWilliam, J., & McCullough, C. M. (2009). A comparison of the hazard perception ability of matched groups of healthy drivers aged 35 to 55, 65 to 74, and 75 to 84 years. *Journal of International Neuropsychological Society, 15*, 799-802.

Horswill, M. S., Waylen, A. E., & Tofield, M. I. (2004). Drivers' ratings of different components of their own driving skill: A greater illusion of superiority for skills that relate to accident involvement. *Journal of Applied Social Psychology, 34*(1), 177–195.

Kantowitz, B. H., Hanowski, R. J., & Kantowitz, S. C. (1997). Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. *Human Factors, 39*(2), 164-176.

Kestin, I. G., Miller, B. R., & Lockhart, C. H. (1988). Auditory alarms during anaesthesia monitoring. *Anesthesiology, 69*(1), 106-109.

Knipling, R. R., Mironer, M., Hendricks, D. L., Tijerina, L., Everson, J., Allen, J. C., et al. (1993). *Assessment of IVHS countermeasures for collision avoidance: rear-end crashes* (No. DOT HS 807 995). Washington, DC: National Highway Traffic Safety Administration.

Kramer, A. F., Cassavaugh, N., Horrey, W. J., Becic, E., & Mayhugh, J. L. (2007). Influence of age and proximity warning devices on collision avoidance in simulated driving. *Human Factors, 49*(5), 935-949.

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors, 43*, 631-640.

Lee, J. D., Gore, B. F., & Campbell, J. L. (1999). Display alternatives for in-vehicle warning and sign information: Message style, location, and modality. *Transportation Human Factors Journal, 1*(4), 347-377.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors, 44*(2), 314-334.

Lee, J. D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics, 35*(10), 1243-1270.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40*, 153-184.

Lee, J. D., & See, K. A. (2004). Trust in technology: Designing for appropriate reliance. *Human Factors, 46*(1), 50-80.

Lees, M. N., Cosman, J. D., Lee, J. D., Rizzo, M., & Fricke, N. (In press). Translating cognitive neuroscience to the driver's operational environment: a neuroergonomics approach. *The American Journal of Psychology*.

Lees, M. N., Cosman, J. D., Lee, J. D., Vecera, S., Jang, M., Dawson, J. D., et al. (2009). *Cross-modal alerts for orienting of attention in attention impaired drivers*. Paper presented at the 17th World Congress on Ergonomics, Beijing, China.

Lees, M. N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics, 50*(8), 1264-1286.

Lees, M. N., & Lee, J. D. (2009). Enhancing safety by augmenting information acquisition in the driving environment. In C. Castro (Ed.), *Human Factors of Visual and Cognitive Performance in Driving*. Boca Raton, FL: CRC Press.

Lerner, N. (1993). Brake reaction times of older and younger drivers. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 1*, 206-210.

Lerner, N. (1994). Giving the older driver enough perception-reaction time. *Experimental Aging Research, 20*, 25-33.

Maltz, M., & Shinar, D. (2004). Imperfect in-vehicle collision avoidance warning systems can aid drivers. *Human Factors, 46*(2), 357-366.

Matthews, M. L., & Moran, A. R. (1986). Age differences in male drivers' perception of accident risk: The role of perceived driving ability. *Accident Analysis & Prevention, 18*(4), 299-313.

Mayhew, D. R., Simpson, H. M., & Pak, A. (2003). Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis & Prevention, 35*, 683-691.

McGwin, G., & Brown, D. B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis and Prevention, 31*(3), 181-198.

McKenna, F. P., & Crick, J. (1994). *Hazard perception in drivers: A methodology for testing and training* (No. Contractor Report 313). Crowthorne, Berkshire: Transportation Research Laboratory: Department of Transport.

McKenna, F. P., Horswill, M. S., & Alexander, J. L. (2006). Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology: Applied, 12*, 1-10.

McKenna, F. P., Stanier, R. A., & Lewis, C. (1991). Factors underlying illusory self-assessment of driving skill in males and females. *Accident Analysis & Prevention, 23*(1), 45-52.

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors, 50*(2), 194-210.

Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors, 46*(2), 196-204.

Meyer, J., & Bitan, Y. (2002). Why better operators receive worse warnings. *Human Factors, 44*(3), 343-353.

Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies, 27*, 527-539.

Myers, R. S., Ball, K., Kalina, T. D., Roth, D. L., & Goode, K. T. (2000). Relation of useful field of view and other screening tests to on-road driving performance. *Perceptual and Motor Skills, 91*, 279-290.

Najm, W. G., Stearns, M. D., Howarth, H., Koopmann, J., & Hitz, J. (2006). *Evaluation of an Automotive Rear-End Collision Avoidance System* (No. DOT HS 810 569): Performed by John A. Volpe National Transportation System Center, Cambridge, MA, Sponsored by National Highway Traffic Safety Administration, Washington, D.C.

Najm, W. G., Wiacek, C. J., & Burgett, A. L. (1998). *Identification of Pre-Crash Scenarios for Estimating the Safety Benefits of Rear-End Collision Avoidance Systems*. Paper presented at the Fifth World Congress on Intelligent Transport Systems, Seoul, Korea, Seoul, Korea.

NHTSA. (2005a). *Automotive Collision Avoidance System Field Operational Test: Final Program Report* (No. DOT HS 809 886). Washington, DC: National Highway Traffic Safety Administration.

NHTSA. (2005b). *Automotive Collision Avoidance System Field Operational Test: Methodology and Results* (No. DOT HS 809 900). Washington, D.C.

NHTSA. (2006). *Traffic safety facts: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System* (No. DOT HS 810 818). Washington, D.C.: National Highway Traffic Safety Administration.

NHTSA. (2009). *2008 Traffic safety Annual Assessement - Highlights*. Washington, D.C.: National Highway Traffic Safety Administration.

Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*, 403-415.

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centred collision-warning systems. *Ergonomics, 40*(3), 390-399.

Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM, 47*(4), 51-55.

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

Pelz, D. C., & Krupat, E. (1974). Caution profile and driving record of undergraduate males. *Accident Analysis and Prevention, 6*, 45-58.

Peterson, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256*, 183-194.

Pollatsek, A., Fisher, D. L., & Pradhan, A. (2006). Identifying and remedying failures of selective attention in younger drivers. *Current Directions in Psychological Science, 15*(5), 255-259.

Pollatsek, A., Narayanaan, V., Pradhan, A., & Fisher, D. L. (2006). Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator. *Human Factors, 48*(3), 447-464.

Preusser, D. F., Williams, A. F., Ferguson, S. A., Ulmer, R. G., & Weinstein, H. B. (1998). Fatal crash risk for older drivers at intersections. *Accident Analysis and Prevention, 30*(2), 151-159.

Quimby, A. R., & Watts, G. R. (1981). *Human Factors and Driving Performance*. Crowthorne: Transport and Road Research Laboratory.

Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine, 336*(7), 453-458.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*(1), 95-112.

Rizzo, M., Anderson, S. W., Dawson, J., Myers, R. S., & Ball, K. (2000). Visual attention impairments in Alzheimer's disease. *Neurology, 54*, 1954-1959.

Roenker, D. L., Cissell, G. M., Ball, K., Wadley, V. G., & Edwards, J. D. (2003). Speed-of-processing and driving simulator training result in improved driving performance. *Human Factors, 45*(2), 218-233.

Sanchez, J., Fisk, A. D., & Rogers, W. A. (2004). *Reliability and age-related effects on trust and reliance of a decision support aid*. Paper presented at the Human Factors and Ergonomics Society 48th Annual Meeting.

Sanders, A. F. (1970). Some aspects of the selective process in the functional field of view. *Ergonomics, 13*, 101-117.

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human Factors, 50*(2), 264-275.

Seagull, F. J., & Sanderson, P. M. (2001). Anesthesia alarms in context: An observational study. *Human Factors, 43*(1), 66-78.

Sekuler, A. B., & Bennett, P. J. (2000). Effects of aging on the useful field of view. *Experimental Aging Research, 26*, 103-120.

Shinar, D. (2000). Fleet study evaluation of an advance brake warning system. *Human Factors, 42*(3), 482-489.

Shinar, D., Rotenberg, E., & Cohen, T. (1997). Crash reduction with an advance brake warning system: A digital simulation. *Human Factors, 39*(2), 296-302.

Smith, K., & Kallhammer, J. (2010). Driver acceptance of false alarms to simulated encroachment. *Human factors, 52*(3), 466-476.

Sorkin, R. D. (1988). Why are people turning off our alarms? *Journal of the Acoustical Society of America, 84*(3), 1107-1108.

Sorkin, R. D., & Woods, D. D. (1985). Systems with human monitors: A signal detection analysis. *Human-Computer Interaction, 1*, 49-75.

Sullivan, J. M., Tsimhoni, O., & Bogard, S. (2008). Warning reliability and driver performance in naturalistic driving. *Human Factors, 50*, 845-852.

Tan, A., & Lerner, N. (1995). *Multiple attribute evaluation of auditory warning signals for in-vehicle crash warning systems* (No. DOT HS 808 535). Washington, DC: National Highway Transportation Safety Administration.

Tränkle, U., Gelau, C., & Metker, T. (1990). Risk perception and age-specific accidents of young drivers. *Accident Analysis & Prevention, 22*(2), 119-125.

Waylen, A. E., Horswill, M. S., Alexander, J. L., & McKenna, F. P. (2004). Do expert drivers have a reduced illusion of superiority? *Transportation Research Part F-Traffic Psychology and Behaviour, 7*(4-5), 323-331.

Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: a synthesis of the literature. *Theoretical Issues in Ergonomics Science, 8*(3), 201-212.

Woods, D. D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics, 38*(11), 2371-2393.

Xiao, Y., & Seagull, F. J. (1999). *An analysis of problems with auditory alarms: Defining the roles of alarms in process monitoring tasks*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting.

Zabyshny, A. A., & Ragland, D. R. P. (2003). *False Alarms and Human-Machine Warning Systems* (No. UCB-TSC-RR-2003-07): UC Berkeley Traffic Safety Center.

Zador, P. L., Krawchuk, S. A., & Voas, R. B. (2000). *Final Report -- Automotive Collision Avoidance (ACAS) Program* (No. DOT HS 809 080). Washington, D.C.: National Highway Traffic Safety Administration.

APPENDIX A.

SCREENING FOR HAZARD AND ALARM EVALUATION

This study will take place in either the Cognitive Systems Lab (CSL) at the University of Iowa or at the University of Iowa Hospital and clinics. There are no known risks associated with this study. You will be viewing a series of brief video clips taken from the driver's perspective on various road types (e.g., rural, urban, highway). While watching the video, you will be asked to identify roadway hazards. Some drivers will watch the video clips with a driver support system that warns the driver about hazards using an auditory tone. Over the course of the experiment you will be asked to fill out a number of questionnaires regarding your driving skills, and your opinions about the videos you watch. Drivers who receive cues will be asked to complete questionnaires about how they perceive the warnings.

The information gathered will be used to answer academic research questions regarding how different drivers identify hazards, and to examine which alerting signals should be placed in cars to improve driver safety. All data obtained are for research purposes only and will remain confidential. Names will not be associated with the data in any way.

The visit will last approximately 1 to 2.4 hours total. We will compensate you $10 per hour up to a possible $24 maximum. You are free to cease participation at any time during the study without risk of penalty.

Do you think that you are interested in participating? ☐ Yes ☐ No

Now I need to ask you a series of questions to verify your eligibility for the study.

1. Is English your first language? ☐ Yes ☐ No
   [*If no, disqualify*]

2. Do you have normal or corrected-to-normal vision? ☐ Yes ☐ No
   [*If no, disqualify*]

3. Do you have a current and valid driver's license? ☐ Yes ☐ No
   [*If no, disqualify*]

4. DOB: (mm/dd/yyyy) __ __ / __ __ / __ __ __ __ Age: ______
   [*Must be between 18-20, between 35 and 50, or between 65 and 98, If no, disqualify*]

5. How many miles do you drive per week on average? _______________
   [For drivers 18-20, miles/week <=100. For drivers 35-50 miles >=100. *If not disqualify]*

6. Have you had any crashes or moving violations? ☐ Yes ☐ No
    - If yes, how many? ____________
    - Specific crashes:

   [For drivers 35-50, # of crashes and moving violations/# years driving <0.5, *if not disqualify*]

Medication / Medical Questions

Medical conditions and some prescription medications can affect driving safety. Before we can determine if you are eligible to participate in this study, we would like to ask you some questions about medical conditions you might have and about any medications you may be taking. Your answers to the following questions are confidential and are used only to determine your eligibility to participate in this study. If you do not qualify for this study, this information will be shredded.

1. What medications are you currently taking? *(Exclude individuals who take prescribed serotonin re-uptake inhibitors (e.g., Lexapro, Zoloft, Paxil, Celexa, Effexor, etc.), lithium carbonate (e.g., Eskalith, Lithane, Lithobid, Cibalith-S, etc.), or other medications treating psychiatric illness unless these are in remission).*

2. **Medical Conditions**: Now I am going to ask you about any medical conditions that you may have. Please answer "Yes", "No", "Don't know", or "In remission" to the following questions. Has a doctor or nurse ever told you that you have…

	Yes	No	Don't Know	In Remission
Head trauma (traumatic brain injury – open or closed head wounds with loss of consciousness)				
Brain Tumors				
Neurodegenerative disease (e.g., Dementia, Alzheimer's Disease, Parkinson's Disease)				
Sleep disorder (e.g. narcolepsy, sleep apnea)				
Epilepsy				
Depression				
Anxiety				
Schizophrenia				
Alcohol or substance abuse problems (e.g., substance abuse may include abuse of stimulants, narcotics, or other illegal substances)				

[If a person answers yes to any of the above questions, inform the person that he/she is not eligible to participate]

Would you be interested in allowing us to keep your name, age, address, and phone number in a contact file so that other investigators can contact you for future research studies.

☐ Yes ☐ No

[If prospective participant indicates "No" to request to keep their name and contact information for future research studies, shred screening sheet].

[If prospective participant indicates "Yes" to request to keep their name and contact information for future research studies, collect their contact information and shred top sheet of this form].

[If prospective participant is eligible for this study schedule a time for him/her to participate. If not eligible end contact]

APPENDIX B.

DEMOGRAPHIC QUESTIONNAIRE

Date of Testing __ __ / __ __ / __ __ __ __

PART I: INFORMED CONSENT AND DEMOGRAPHIC INFORMATION

1. **Name**

   Last: ________________________ First:____________ Middle Initial:______

2. **Address**

   Address Line 1:____

   Address Line 2:____

   City:________________ State: _________________ Zip Code: __ __ __ __ __

3. **Contact Information**

   Primary Phone Number: __ __ __ - __ __ __ - __ __ __ __

   ☐1 Home phone ☐3 Cell phone

   ☐2 Work phone

   Email:___________________________@_______________________

4. **DOB** __ __ / __ __ / __ __ __ __ (*mm/dd/yyyy*)

5. __ __ **Age (Years)**

6. **Handedness** ☐1 Right ☐2 Left ☐3 Mixed

7. **Gender:** ☐1 Male ☐2 Female

8. **Vision: Normal** ☐1 Yes ☐0 No

9. **Corrected to Normal: with Glasses?** ☐1 Yes ☐0 No

10. **Corrected to Normal: with Contacts?** ☐1 Yes ☐0 No

11. **Education:** What is the highest grade of school or level of education that you <u>completed</u>?

☐00 Did not go to school
☐01 Grade 1
☐02 Grade 2
☐03 Grade 3
☐04 Grade 4
☐05 Grade 5
☐06 Grade 6
☐07 Grade 7
☐08 Grade 8
☐09 Grade 9
☐10 Grade 10
☐11 Grade 11
☐12 Grade 12/GED
☐13 Vocational training / some college
☐14 Associate degree
☐16 College graduate (BA or BS degree)
☐17 Some professional school (after college)
☐18 Master's degree
☐20 Doctoral degree (PhD, MD, DVM, DDS, JD, etc.)

12. What is your current marital status?

☐1 Married
☐2 Living as married
☐3 Separated
☐4 Divorced
☐5 Widowed
☐6 Single, never married

13. What race do you consider yourself?

☐1 White/Caucasian
☐2 Black/African American
☐3 Asian
☐4 Native Hawaiian/Pacific Islander
☐5 American Indian/Alaskan Native
☐6 Biracial: ____________________
☐7 Other: ____________________
☐8 Don't know

14. Do you have any Hispanic or Latino background? ☐ 1 Yes ☐0 No

15. Driving Information:

15a. Do you have a current and valid driver’s license? ☐ 1 Yes ☐0 No

If yes: Exp. __ __ / __ __ / __ __ __ __

15b. How many years have you been driving? __ __ (*Years*)
If necessary, prompt with the following options: (check the answer)

☐0 Never driven
☐1 1 – 5 years
☐2 6 – 10 years
☐3 11 – 15 years
☐4 > 15 years
☐5 Don’t know

15c. How many miles per week do you drive? __ __ __ (*miles*)

If necessary, prompt with the following options: (check the answer)

☐0 Don’t drive currently
☐1 1 – 50 miles / week
☐2 51- 100 miles / week
☐3 101 - 150 miles / week
☐4 151 - 200 miles / week
☐5 201 - 250 miles / week
☐6 > 250 miles / week
☐7 Don’t know

15d. How many days of the week do you drive?

☐0 0 days / week
☐1 1 day / week
☐2 2 days / week
☐3 3 days / week
☐4 4 days / week
☐5 5 days / week
☐6 6 days / week
☐7 7 days / week
☐8 Don’t know

15e. When was the last time you drove a car, excluding today?

☐0 yesterday
☐1 within the past week
☐2 within the past month
☐3 within the past year
☐4 more than one year ago
☐5 Don’t know/don’t Remember

15f. Have you had any crashes or moving violations within the last 5 years?

☐1 Yes ☐0 No

If yes: how many crashes? ____________

If yes: how many moving violations? ____________

Type of moving violations? ___________

APPENDIX C.

INTERPERSONAL TRUST QUESTIONNAIRE

Please mark the degree to which you agree with the following statements:

	Strongly Agree	Mildly Agree	Agree and Disagree Equally	Mildly Disagree	Strongly Disagree
Hypocrisy is on the increase in our society.	○	○	○	○	○
In dealing with strangers one is better off to be cautious until they have provided evidence that they are trustworthy.	○	○	○	○	○
This country has a dark future unless we can attract better people into politics.	○	○	○	○	○
Fear and social disgrace or punishment rather than conscience prevents most people from breaking the law.	○	○	○	○	○
Using the honor system of not having a teacher present during exams would probably result in increased cheating.	○	○	○	○	○
Parents usually can be relied on to keep their promises.	○	○	○	○	○
The United Nations will never be an effective force in keeping world peace.	○	○	○	○	○
The judiciary is a place where we can all get unbiased treatment.	○	○	○	○	○
Most people would be horrified if they knew how much news that the public hears and sees is distorted.	○	○	○	○	○
It is safe to believe that in spite of what people say most people are primarily interested in their own welfare.	○	○	○	○	○
Even though we have reports in newspaper, radio, and TV, it is hard to get objective accounts of public events.	○	○	○	○	○
The future seems very promising.	○	○	○	○	○
If we really knew what was going on in international politics, the public would have reason to be more frightened than they now seem to be.	○	○	○	○	○
Most elected officials are really sincere in their campaign promises.	○	○	○	○	○
Many major national sports contests are fixed in one way or another.	○	○	○	○	○

	Strongly Agree	Mildly Agree	Agree and Disagree Equally	Mildly Disagree	Strongly Disagree
Most experts can be relied upon to tell us the truth about limits of their knowledge.	○	○	○	○	○
Most parents can be relied upon to carry out their threats of punishment.	○	○	○	○	○
Most people can be counted on to do what they say they will do.	○	○	○	○	○
In these competitive times one has to be alert or someone is likely to take advantage of you.	○	○	○	○	○
Most idealists are sincere and usually practice what they preach.	○	○	○	○	○
Most salesmen are honest in describing their products.	○	○	○	○	○
Most students in school would not cheat even if they were sure they would get away with it.	○	○	○	○	○
Most repairmen will not overcharge even if they think you are ignorant of their specialty.	○	○	○	○	○
A large share of accident claims are filed against insurance companies are phony.	○	○	○	○	○
Most people answer public opinion polls honestly.	○	○	○	○	○

APPENDIX D.

DRIVING SKILL QUESTIONNAIRE

Please rate your ability on the following

	Very Weak	Weak	Avg	Strong	Very Strong
Fluent driving	○	○	○	○	○
Performance in critical situations	○	○	○	○	○
Perceiving hazards in traffic	○	○	○	○	○
Driving in a strange city	○	○	○	○	○
Paying attention to pedestrians and cyclists	○	○	○	○	○
Driving on a slippery road	○	○	○	○	○
Conforming to traffic rules	○	○	○	○	○
Managing a car through a slide	○	○	○	○	○
Preview of traffic situations	○	○	○	○	○
Driving carefully	○	○	○	○	○
Control of the traffic situation	○	○	○	○	○
Fluent lane changing in heavy traffic	○	○	○	○	○
Fast reactions	○	○	○	○	○
Making firm decisions	○	○	○	○	○
Paying attention to other road users	○	○	○	○	○
Driving fast if necessary	○	○	○	○	○
Driving in the dark	○	○	○	○	○
Controlling the vehicle	○	○	○	○	○
Avoiding competition in traffic	○	○	○	○	○
Keeping sufficient following distance	○	○	○	○	○
Adjusting speed to the present traffic conditions	○	○	○	○	○
Overtaking	○	○	○	○	○
Cleaning the car windows on winter mornings	○	○	○	○	○
Relinquishing one’s rights	○	○	○	○	○
Conforming to the speed limits	○	○	○	○	○
Avoiding unnecessary risks	○	○	○	○	○
Tolerating other drivers’ blunders calmly	○	○	○	○	○
Following the traffic lights carefully	○	○	○	○	○
Parking in legal places only	○	○	○	○	○

APPENDIX E.

HAZARD EVALUATION INSTRUCTIONS

You will be shown video clips of traffic situations filmed from a driver’s point of view. We want you to watch the videos as if you are the driver.

The videos may contain scenarios where a collision (or near collision) between you and another vehicle might occur unless you take evasive action (such as slowing).

While watching the videos, your task is to indicate when you would begin to slow down to avoid a traffic conflict. If a scenario involving a potential conflict occurs that you did not anticipate, then it is better to respond late rather than not at all.

During each video, please press the mouse button as soon as you would initiate a response. If you feel that no traffic conflict exists you should not make a response.

Even if you make a response the video will continue to play.

If you begin to feel nauseous, please let the researcher know immediately.

APPENDIX F.

VIDEO EVALUATION

How hazardous was the situation you just viewed?

	Not At All Hazardous	Slightly Hazardous	Somewhat Hazardous	Moderately Hazardous	Extremely Hazardous
In this situation the alarm was	○	○	○	○	○

What was the hazard of conflict?

How confident are you that you could avoid a collision in situations like the one you just saw?

	Not At All Confident	Slightly Confident	Somewhat Confident	Moderately Confident	Extremely Confident
Based on this situation I am	○	○	○	○	○

APPENDIX G.

INSTRUCTIONS ON FORWARD COLLISION WARNING SYSTEM

(Adapted from NHTSA, 2005b)

FCW becomes active when you are travelling at speeds greater than 25 mph.

FCW provides audio warnings if you rapidly approach a slow or stopped object. FCW will also warn if you are following another vehicle too closely. FCW is only a warning system and is alerting you to the potential for a collision. FCW does not decelerate the vehicle.

You may receive an alarm when another vehicle cuts in front of you. A car slowing in order to make a right or left hand turn may also prompt an alarm. There are infrequent instances when the FCW system may present an alarm when no hazard actually exists. Examples of road side objects that might produce an alarm are mailboxes, sign posts, light poles and guard rails. In particular, these items may produce an alarm if they are located close to the road's edge or straight ahead of you as you approach a curve.

You as the driver, must be prepared to apply the brakes whenever you are presented with the imminent warning.

APPENDIX H.

ALARM EVALUATION QUESTIONNAIRE

How hazardous was the situation you just viewed?

	Not At All Hazardous	Slightly Hazardous	Somewhat Hazardous	Moderately Hazardous	Extremely Hazardous
In this situation the alarm was	○	○	○	○	○

What was the hazard of conflict?

How easy was it to understand why the alarm occurred in this situation?

	Not At All Easy to Understand	Slightly Easy to Understand	Somewhat Easy to Understand	Moderately Easy to Understand	Extremely Easy to Understand
In this situation the alarm was	○	○	○	○	○

Why was the alarm given?

How useful was the alarm in this situation?

	Not At All Useful	Slightly Useful	Somewhat Useful	Moderately Useful	Extremely Useful
In this situation the alarm was	○	○	○	○	○

How appropriate was the alarm in this situation?

	Not At All Appropriate	Slightly Appropriate	Somewhat Appropriate	Moderately Appropriate	Extremely Appropriate
In this situation the alarm was	○	○	○	○	○

How would you rate the timing of the alarm?

	Early	Correct	Late	N/A
In this situation the timing of the alarm was	○	○	○	○

Based on its behaviour in this video, how trustworthy is the system?

	Not At All Trustworthy	Slightly Trustworthy	Somewhat Trustworthy	Moderately Trustworthy	Extremely Trustworthy
Based on this situation the system is	○	○	○	○	○

How confident are you that you could avoid a collision in situations like the one you just saw?

	Not At All Confident	Slightly Confident	Somewhat Confident	Moderately Confident	Extremely Confident
Based on this situation I am	○	○	○	○	○