Individual Differences in Intramodal and Crossmodal Inattentional Insensitivity and The Design of In-Vehicle Alert Systems

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Bridget A. Lewis
Master of Arts
George Mason University, 2012
Bachelor of Sciences
Virginia Polytechnic Institute and State University, 2010

Director: Carryl L. Baldwin, Professor
Department of Psychology

Fall Semester 2017
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I would like to thank my family, Mom, Dad, Mike and Aidan, for their constant, unwavering support, and unconditional love. I would like to thank my teachers, both official and unofficial, and particularly my mentor, Carryl, for their advice and their encouragement. I would like to acknowledge the work of my labmates, without whom much of this research would never have happened, for their direct support of this research and indirect support through their help with brainstorming, reassurance, and willingness to help lug computers all over the building. I would finally like to thank Devon. For pushing me, supporting me, helping me, and believing in me.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

ADAS	Advanced Driver Assistance Systems
BRT	Brake Response Time
CAMP	Crash Avoidance Metric Partnership
CIS	Crossmodal Inattentional Insensitivity
CS	Critical Stimulus/Signal
DVI	Driver Vehicle Interface
EMRT	Emergency Maneuver Response Time
FARS	Fatal Accident Reporting System
FCW	Forward Collision Warning
GMU	George Mason University
IBI	Interburst Interval
IIS	Intramodal Inattentional Insensitivity
IPI	Interpulse Interval
ISI	Interstimulus Interval
LBFTS	Looked But Failed to See
LED	Light Emitting Diode
M	Mean
NHTSA	National Highway Traffic Safety Administration
OSpan	Operational Span Task
PEBL	Psychology Experiment Building Language Program
RGB	Red Green Blue Value
RSAP	Rapid Serial Auditory Presentation
RTI	Realtime Technologies, Inc.
SAS	Simulator Adaptation Syndrome
SD	Standard Deviation
WMC	Working Memory Capacity

ABSTRACT

INDIVIDUAL DIFFERENCES IN INTRAMODAL AND CROSSMODAL INATTENTIONAL INSENSITIVITY AND THE DESIGN OF IN-VEHICLE ALERT SYSTEMS

Bridget A. Lewis, Ph.D

George Mason University, 2017

Dissertation Director: Dr. Carryl L. Baldwin

Around 20% of all automobile crashes in recent years have been linked to driver distraction or inattention. A subset of these crashes, involve “Looked But Failed to See” (LBFTS) incidents in which an otherwise attentive driver completely fails to notice a salient signal. In the best case, this may involve a driver putting on the brakes late because she failed to notice a red light and stopping with her nose into an intersection, causing embarrassment but no harm. But in the worst case, looking but failing to see causes about 6% of all injury and fatality-related crashes per data from the Fatal Accident Reporting System (FARS) maintained by the National Highway Traffic Safety Administration (NHTSA). LBTFS are a type of inattentional blindness, a subset of inattentional insensitivity. Inattentional insensitivity is a blanket term that describes the well-known phenomena of inattentional blindness and inattentional deafness. These phenomena occur when an otherwise salient stimulus is missed during high levels of perceptual load. For example, pilots coming in for a difficult landing in high cross winds miss auditory alarms notifying them of a landing gear failure, or drivers lost at night on an unfamiliar route fail to see a stop sign. It is hypothesized that inattentional insensitivity is integrally tied to an individual's working memory capacity. Previous studies have proposed theoretical accounts involving both single and dual routes for this relationship. A major goal of this dissertation is to determine which of these theoretical explanations best predict patterns of inattentional insensitivity. A second goal of this dissertation is to address methods of ameliorating inattentional insensitivity, regardless of their cause, via the design of effective, multimodal alert systems. Towards the second goal, the first study, details an investigation of multimodal urgency scaling with the goal of determining perceived changes in urgency relative to physical changes in visual, auditory and tactile stimuli. Psychometric functions were obtained for various parameters within each modality based on perceptions of urgency, annoyance and acceptability. Results indicated that auditory stimuli affected the biggest increases in urgency relative to physical changes, but that with increased urgency often came increased annoyance. Visual stimuli were rarely rated as annoying but were also unable to achieve similarly high levels of urgency relative to auditory or tactile stimuli. Tactile stimuli showed the greatest utility (indicating greater urgency changes in relation to annoyance changes). The second study, was designed to validate the psychometric functions established in the first study by examining behavioral responses to warnings designed to be perceived as highly urgent and time critical versus warnings missing key parameters within the context of driving. Specifically, the second study examined the potential for appropriate warnings to eliminate inattentional insensitivity to alerts while

distracted, regardless of why it was occurring. Towards this aim, the second study required participants to drive a simulated course while completing a distracting task and following a lead vehicle. At a pre-set point, the lead vehicle swerved sharply into the left lane to avoid a revealed, stopped car. Participants then received either a "good" warning, one that met all pre-defined criteria, an "edge" warning, one that met only some of the criteria, or no warning. Results indicated that, while crash occurrences were not significantly different, for those who did crash, they crashed at a significantly slower speed.

The final study in this series, sought to examine the effect of working memory capacity (WMC: as measured by OSpan) on various types of inattentional insensitivity. Specifically, inattentional insensitivity was examined for intramodal and crossmodal tasks involving either visual or auditory critical signals. Participants were asked to first complete a computerized version of the OSpan task to evaluate their working memory capacity. They were then assessed for inattentional insensitivity starting in one of four conditions: visual task- visual critical signal, visual task-auditory critical signal, auditory task-auditory critical signal, and auditory task-visual critical signal. In this series, the visual task was a cross arm-length detection task and the auditory task was a rapid serial auditory presentation task. The critical signals (CSs) were either visual pictures of shapes or auditory names of shapes. Results from Study 3 indicate that, although inattentional insensitivity was present in all modality combinations for some proportion of the population, individuals were less likely to miss critical signals when they were in the same modality as the main task (intramodal signals). Results also indicate a significant difference in sensitivity by WMC, where those with medium to high WMC were significantly more likely to notice an intramodal CS than a crossmodal CS, though this effect was not present for those with low WMC levels.

Results from the present series of studies inform the design of in-vehicle alerting systems, and may be particularly perteinent for highly automated or autonomous vehicles. The drivers or , for lack of a better word, operators of these vehicles may well be fatigued, distracted, or otherwise impaired which will increase the need for targeted, highly effective warnings when operator intervention is required.

1 INTRODUCTION

In many situations people fail to notice salient stimuli. In the case of driving, stop signs are missed and people do not notice a warning light or auditory alarm while their attention is focused on another task. In fact, according to the Fatal Accident Reporting System (FARS) maintained by the National Highway Traffic Safety Administration (NHTSA), in 2013 around 20% of all injury-related crashes were due to some type of driver distraction. Furthermore, between 5 and 6 percent of all fatalities since 2009 (the first year in which NHTSA FARS data includes distraction related qualifiers) have been related to inattention. Inattention, in the FARS database, includes “looked but did not see” accidents also known as Looked But Failed to See (LBFTS) accidents, in which the driver was paying attention to the driving task, but failed to notice a roadway hazard. These accidents make up about 5% of all distraction-related injuries and fatalities and about 1% of all crashes involving injury or fatalities. This effect isn’t limited to driving. Airline pilots have been known to miss highly salient auditory alerts when attempting to complete a difficult landing (Dehais et al., 2014), police officers chasing a suspect may miss on-going assaults (Chabris, Weinberger, Fontaine, & Simons, 2011), and even a unicycling clown can be missed in broad daylight (Hyman, Boss, Wise, McKenzie, & Caggiano, 2010). Despite the relative celebrity of this effect, and a large breadth of research investigating when or under what circumstances this effect occurs, there has been no clear explanation for why or to whom it occurs. Understanding when people are more likely to fail to notice salient stimuli in real world contexts and which individuals or groups are more likely to have increased susceptibility is critical to maintaining safety.

There are various, sometimes opposing, theories that may explain why a person might fail to notice a salient signal (discussed in further detail below). Researchers have theorized that individual differences may have something to do with this failure to notice, and one of the most well-researched individual differences (with respect to inattentional insensitivity paradigms) is working memory capacity (WMC). Findings related to WMC and inattentional insensitivity have been inconclusive, however recently it has been theorized that there may be a dual-route model, a U-shaped function, for inattentional insensitivity by WMC as proposed by Hannon and Richards and colleagues (Hannon & Richards, 2010; Richards, Hannon, & Derakshan, 2010). Specifically, they theorize that individuals with low WMC fail to notice salient stimuli when under high perceptual load because they are unable to process additional stimuli which are unrelated to the central task, but that individuals with high WMC fail to notice because they over-inhibit attention to distractors. This theory leads to the prediction that individuals in the middle range of WMC should have the lowest rates of inattentional insensitivity. Further, the dual route theory leads to the prediction that, possibly, low WMC individuals would be less inattentionally insensitive to crossmodal stimuli but high WMC individuals might still be able to inhibit any irrelevant signals, including bimodal.

Signals must be appropriately designed such that they both capture attention and unambiguously inform users to their purpose, but even well-designed signals can be missed. The purpose of this dissertation is to closely examine inattentional insensitivity, to investigate appropriate warning design for mitigation of inattention and to examine the effect of individual differences on inattentional insensitivity in order to better understand the relationship between warning design and missed signals. Before describing a series of studies, the key theoretical concepts are defined, and relevant literature is reviewed.

1.1 Defining Inattentional Insensitivity

Inattentional Insensitivity constitutes a blanket term used to describe all phenomena whereby individuals fail to perceive would-be salient cues as a direct result of inattention. In particular, people fail to notice a salient cue when their attention is directed to a high demand perceptual task (Dattel et al., 2013; Mack & Rock, 1998). Simons (2007) gives four specific criteria for classifying a failure of awareness as inattentional blindness which can be modified to apply to all inattentional insensitivity: 1) that observers must fail to notice an object or event; 2) that the object must be fully perceptible and readily noticed when it is being waited for (looked for, listened for, felt for, etc.) specifically; 3) that the failure to notice the object or event results from a lack of engagement of attention to the object or event, not from physical aspects of the object or event itself; and, 4) that the object or event must be unexpected. These criteria can be adjusted to fit any inattentional insensitivity paradigm by adjusting the modality of the object or event. Inattentional insensitivity has been studied in a variety of contexts under a variety of names.

Research conducted to date has been inconsistent in both methodology and in findings and conclusions. There is some evidence indicating that inattentional insensitivity may be related to critical stimulus salience, primary task difficulty, expertise on primary task, age (broadly), sensory and processing abilities, working memory capacity and cognitive control. However, these findings are often conflicting, and comparisons cannot readily be made between results as methodologies vary so widely and often fail to account for individual differences.

This introduction will cover early accounts of inattentional insensitivity including selective listening tasks, field observations and the introduction of the term "inattentional blindness" as this preliminary work is essential to the understanding of the phenomenon. It will then summarize findings for both intramodal and crossmodal inattentional insensitivity as they have been investigated to date, specifying where there are gaps or inconsistencies. It will also cover findings on behavioral and societal correlates of inattentional insensitivity, and manipulations of emotional salience as they relate to inattentional insensitivity. It will continue to include overviews of signal and display design to mitigate the effect of inattentional insensitivity in the context of vehicle operations. Finally, it will include overviews of individual differences that have been found to be related to inattentional insensitivity, including the relationship between working memory capacity, executive function and inattentional insensitivity. These individual differences findings are of particular importance in that they help us to understand where and when and to whom insensitivity is more likely to occur, despite the inclusion of alerting assistance systems. Some of the earliest accounts of what we now consider inattentional blindness included studies done by Neisser, Becklen and colleagues in the late 1970's on what they termed "selective looking" (Becklen & Cervone, 1983; Neisser, 1979; Neisser & Becklen, 1975).

1.2 Selective Looking and Attentional Capture

Neisser and colleagues attempted to create a visual version of the selective listening and selective reading paradigms being widely used in auditory and verbal processing studies in the middle of the 20th century. Selective listening studies involve the listener attending to one stream of sound while ignoring a second stream and, similarly, selective reading studies involve reading one line of a story while ignoring a second line. Typically, findings of these studies indicate that subjects perform well when asked to recall content from the attended stream but do very poorly when asked to recall content in the unattended stream, with a few exceptions. The early work of Neisser and colleagues is particularly important as it established the groundwork for inattentional insensitivity paradigms. Neisser and Becklen (1975) were specifically interested in whether the phenomena observed in selective listening and selective reading studies would be observed in selective looking studies. They used a paradigm first introduced in the late 1960's by Kolers (1969) in which subjects wearing a half-silvered mirror headgear saw superimposed images of what was in front of them and what was behind them and were asked to monitor only one stream at once. Neisser and Becklen used a similar paradigm to allow subjects to view superimposed images of two "games" and asked them to monitor target events. The games used included a "hand game", where two players stand palm to palm and the player with his hands on the bottom attempts to slap the top of the other player's hands, and a "ball game", where subjects monitored ball passes by three players moving irregularly in the view frame. Target events in the hand game included attack moves by the player with his hands on the bottom but not feints (or instances in which the player jerks his hands but does not actually attack), and target events in the ball game included passes by the players but not dribbles or fake throws. Importantly, "odd" events were included in both games. Odd events in the hand game included the players stopping in the middle of the game and shaking hands then resuming play, or throwing a ball from one player to the other then resuming play, and odd events in the ball game included the disappearance of the ball for a short period of time with the players playing as normal until the ball was brought back into the game, or an exchange where male players left one by one and were replaced by female players for a short time, then reentered the game, replacing the female players. These scenes were either displayed to participants superimposed over each other using a mirror system or displayed such that each eye saw a different video (**Error! Not a valid bookmark self-reference.**).

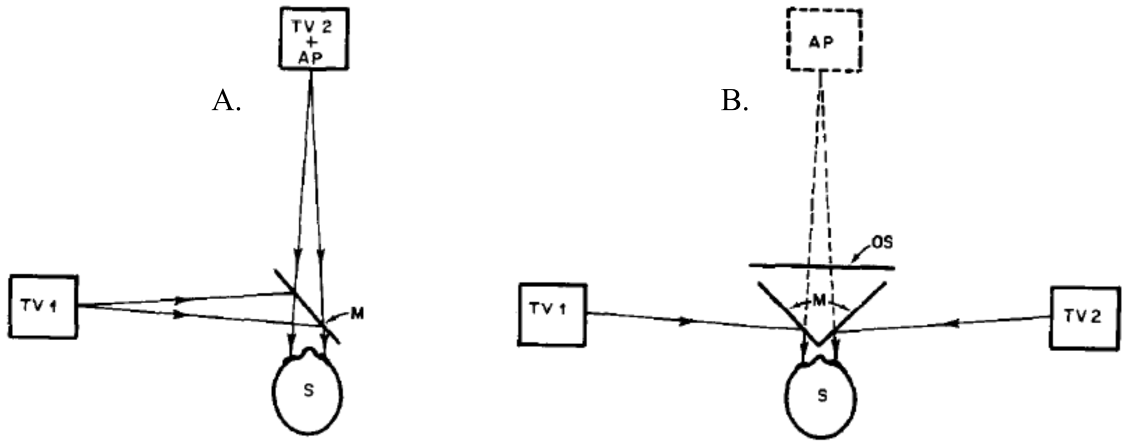

Figure 1. Schematic view of optical arrangement from Neisser & Becklen (1975). A. Binocular condition. B. Dichoptic condition. Abbreviations: S= Subject; TV = Television monitor; 1 = Ball game, 2 = Hand game; M= Half-silvered mirror; OS = Opaque screen; AP = Apparent position of superimposed images.

After each of ten trials subjects were asked about their strategies and anything else they might want to report, then at the end of the experimental session subjects were asked more extensive questions including specifically the four possible odd events and whether subjects had noticed them.

While subjects had very little difficulty following target events in attended streams, they very rarely noticed anything out of the ordinary in the unattended stream. Only one subject (out of 24 included in the study) noticed and spontaneously reported the handshake in the hand game while monitoring the ball game (with three more reporting it in the post experimental questioning). No subject noticed the ball disappearance in the ball game when watching the hand game. Six subjects in total noticed the ball throw in the hand game when monitoring the ball game, four of whom were the same subjects that had already reported noticing the handshake, and five subjects noticed the male-female change in the ball game when monitoring the hand game. Half of the subjects in the experiment never noticed anything and the half that did notice always gave partial answers, never being fully aware of all aspects of an odd event. Neisser and Becklen also point out that pilot subjects viewing only one video passively always noticed odd events, implying that the events are, in-and-of themselves, very noticeable. Interestingly, one subject following the ball game did not report seeing the ball throw in the hand game but did respond as though to a target ball pass indicating that subconsciously, he did perceive the event. However, when questioned he insisted that he had not actually seen the ball being thrown (Neisser & Becklen, 1975). These findings led Neisser & Becklen to conclude that, rather than ignoring or suppressing irrelevant streams, subjects are merely failing to allocate any attention to them. In their words "One event is perceived because the relevant information is being picked up and used; other information is not picked up in the first place, and consequently not used" (p. 494). This study, while interesting in itself, did little to explain the theory behind why some subjects did not notice critical events. However, they spawned a new line of research investigating the role of attention, and more importantly, inattention.

One early theory of inattentional insensitivity was that people notice the odd event but that they forget by the end of the task and therefore are no longer aware of an event they originally perceived. Wolfe (1999) referred to this theoretical account as inattentional amnesia. Wolfe's hypothesis proposes that unexpected events are consciously perceived but that they are then immediately forgotten when they leave the visual field, meaning that non-noticing in these paradigms is not a perceptual failure but a failure of memory.

Becklen & Cervone (1983) systematically examined the inattentional amnesia hypothesis by manipulating aspects of the task in an attempt to discover if the non-reporting of a critical event was actually just the participants' forgetting due to the complexity of the main task and the time between the event and the end of the trial. They found no difference in noticing of an unexpected woman carrying an umbrella through two superimposed passing games between groups that saw a full 1-minute clip and those who were asked about what they saw directly after the woman exited the screen. This lack of a difference in rates of noticing was taken by Becklen & Cervone to indicate that the notion of forgetting the critical stimulus was likely inaccurate. Although the critical stimulus may have been implicitly noticed, it was not explicitly attended to and therefore not processed into working memory by those subjects who exhibited inattentional insensitivity.

Further, Becklen and Cervone (1983) found that even if the video was stopped while the woman was still on the screen, many subjects still reported nothing out of the ordinary. Simons & Chabris (1999) point out that in recordings of subjects completing the gorilla task those who report having noticed the event typically smile or laugh when the gorilla comes on screen, but non-noticers show no signs of anything being different in the task.

1.3 "Inattentional Blindness"

These early studies by Neissner, Becklen and colleagues were soon followed by many studies investigating these types of attentional paradigms. The name "selective looking" was replaced as various studies tried to pick out more specific issues and processes. In the late 1990s, Mack & Rock, (1998) described in detail a new method for systematically investigating the phenomenon that they called "inattentional blindness". They specifically point out that the difference between inattentional blindness studies and those which attempted to study attention and attentional capture, is that distraction and visual search paradigms both involve telling subjects that there may be a target and sometimes what that target actually is. In other words, subjects will inevitably direct some attention to the potential targets if they have been told to expect one. This is true in both divided attention paradigms and in parallel search paradigms. The novelty in their "new method" of lab-based design was that their paradigm ensured that subjects would be looking *at* the critical signal, the unexpected target, but would not be looking *for* the critical signal. Their task involved the presentation of a fixation cross followed by a large cross centered at fixation. The cross had one long and one short arm followed by a mask. Subjects in this task were asked to report the long arm on each trial. In a later trial, a critical stimulus (CS) appears. In this type of basic study, the CS is a small but noticeable dot, noticed nearly all the time by observers not doing the cross task. Other researchers have extended this paradigm across modalities, using auditory and, more recently, tactile main tasks and critical stimuli.

1.4 Multimodal Inattentional Insensitivity

Mack and Rock (1998) further hypothesize that inattentional blindness is just a visual manifestation of inattentional insensitivity, which they describe as a failure to perceive as a result of inattention. This varies from Wolfe's (1999) hypothesis in that Wolfe believed that the CS is perceived and attended to but is never fully encoded into working memory whereas Mack and Rock believe that the CS does not even reach conscious perception. They hypothesize that this effect should be present in all sensory systems, not just the visual domain. Indeed, it has since been observed that the phenomenon exists in the auditory domain, termed inattentional deafness (Macdonald & Lavie, 2011). Investigations of inattentional deafness have taken two separate routes. The first is deafness observed under conditions of high visual load, and second is deafness observed under conditions of high auditory load. Interestingly, this allows us to separate studies of inattentional insensitivity into intramodal inattentional insensitivity - where insensitivity occurs in the same modality as the perceptually loading task, and crossmodal inattentional insensitivity - where insensitivity occurs in a modality different to the perceptually loading task.

Perceptual load theory (Lavie & Tsal, 1994) posits that in cases of high perceptual load, there will not be enough attention left for task irrelevant stimuli. However, multiple resource theory (Wickens, 2002) predicts that tasks in separate modalities should be completed more efficiently as long as they don't overlap in any central processing areas at the same time. These theories should converge in their predictions for intramodal and crossmodal inattentional insensitivity in that intramodal paradigms should show higher levels of insensitivity than crossmodal paradigms. However, this is not necessarily the case as varying levels of inattentional insensitivity have been found in all paradigms (see Table 1).

Table 1. Intramodal and crossmodal inattentional insensitivity across studies (letters following years indicate the order of the studies in each paper)

Crossmodal Inattentional Blindness			
Study	Sample Size	Sample Age	Inattentionally Insensitive
Beanland, Allen & Pammer (2011a)	25	20.9	28%
Beanland, Allen & Pammer (2011c)	25	20.9	56%
Hyman et al (2010a)	24	Students	75%
Hyman et al (2010b)	28	Students	39%
Crossmodal Inattentional Deafness			
Dehais et al (2012)	14	29.9	43%
Macdonald & Lavie (2011a)	49	22	75%
Macdonald & Lavie (2011b)	39	21.1	79%
Macdonald & Lavie (2011c)	39	22.2	44%
Intramodal Inattentional Blindness			
Arndt et al (2006)	30	32.6	90%
Beanland, Allen & Pammer (2011b)	25	20.9	80%
Beanland, Allen & Pammer (2011d)	25	20.9	52%
Beanland, Allen & Pammer (2011e)	10	21.3	40%
Becklen & Cervone (1983)	85	Undergraduate	65%
Bredemeier & Simons (2012a)	134	19	29%
Bredemeier & Simons (2012b)	207	19.5	73%
Chabris et al (2011a)	15	Students	44%
Chabris et al (2011b)	33	Students	58%
Chabris et al (2011c)	25	Students	28%
Chabris et al (2011d)	20	Students	65%
Clifasefi et al (2006a)	47	21-35	82%
Clifasefi et al (2006b)	47	21-35	50%
Dattel et al (2013b)	36	Students	55%
Drew, Vo & Wolve (2013a)	24	48	92%
Drew, Vo & Wolve (2013b)	25	33.7	100%
Graham & Burke (2011a)	35	61-81	45%
Graham & Burke (2011b)	26	61-81	90%
Graham & Burke (2011d)	31	17-22	40%
Hannon& Richards (2010)	77	31	50%
Kennedy & Bliss (2014)	44	20.84	21%
New & German (2014b)	252	19	47%
O'Shea & Fieo (2014)	36	69.54	74%
Remington, Cartwright-Finch & Lavie (2014a)	40	12	50%
Remington, Cartwright-Finch & Lavie (2014b)	32	14	50%
Remington, Cartwright-Finch & Lavie (2014c)	40	7 y 11 m	83%
Remington, Cartwright-Finch & Lavie (2014d)	44	9 y 11 m	77%
Remington, Cartwright-Finch & Lavie (2014e)	32	30	16%
Seegmiller, Watson & Strayer (2011)	197	18-35	42%
Simons & Chabris (1999)	192	Undergraduate	46%
Simons & Jensen (2009a)	43	Undergraduate	29%
Simons & Jensen (2009b)	82	Undergraduate	34%
Simons & Jensen (2009c)	43	Undergraduate	68%
Simons & Jensen (2009d)	43	Undergraduate	51%
Intramodal Inattentional Deafness			
Dalton & Fraenkel (2012a)	20	20	10%
Dalton & Fraenkel (2012b)	20	20	70%
Dalton & Fraenkel (2012c)	20	20	35%
Dalton & Fraenkel (2012d)	20	20	55%

Some researchers have theorized that observed differences in levels of inattentional insensitivity may have to do with the observers' use of an attentional set (Folk, Remington, & Johnston, 1992; Koivisto & Revonsuo, 2008). An attentional set includes a range of features relevant to the primary task, for example, in the instance of the well-known gorilla video, an observer asked to count passes by the white team would have adopted an attentional set including the ball and the white-shirted players, an example of a strong attentional set. Beanland, Allen, & Palmer (2011) suggested that inattentional blindness may become less pronounced when observers are forced to adopt a weaker (or broader) attentional set. In their study, participants either completed a low or a high load visual task (tracking the letters L and T as they bounced slowly or quickly around a screen, respectively) or completed a low load visual task (passively viewing the screen) while simultaneously listening to music which either had or did not have embedded tones (to which high auditory load participants were also asked to respond verbally). The critical stimulus in this case was the appearance and subsequent screen crossing of the letter A (see **Error! Reference source not found.** 2).

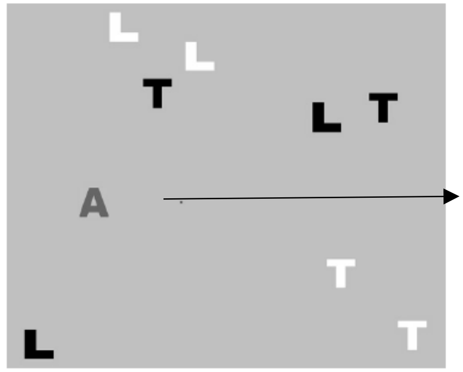

Figure 2. Inattentional blindness paradigm used by Beanland, Allen & Pammer (2011).

Beanland, Allen and Pammer (2011) found that rates of inattentional blindness were lowest in the high auditory load condition. Participants in the most demanding perceptual condition actually noticed the critical stimulus the most. The authors suggest that this may be due to the forced adoption of a weak attentional set (e.g. follow letters and music) thereby un-focusing attention as opposed to other conditions in which participants might have adopted a relatively strong attentional set (e.g. follow white letters, ignore everything else) which would imply highly focused attention (Beanland et al., 2011). It is particularly important to point out that studies investigating various forms of inattentional insensitivity have been undertaken through dramatically differing methods ranging from computer-based laboratory tasks to realistic or real-world tasks and differences in levels of inattentional insensitivity may a be attributed to variation in the paradigms and methods used to study insensitivity.

1.5 Inattentional Insensitivity Paradigms

Inattentional insensitivity paradigms can be grouped into one of three types. The first type of tasks is a computer based paradigm, in which the main task typically involves the following of a visual or auditory stream of simple stimuli and the CS is typically some item that is dissimilar to stimuli used in the main task. These studies have shown ranges of inattentional insensitivity between 16% (Remington, Cartwright-Finch, & Lavie, 2014) for adults on a cross distinction task and 83% for a younger sample on the same task (see also Beanland et al., 2011; Bredemeier & Simons, 2012; Hannon & Richards, 2010; Macdonald & Lavie, 2011; Mack & Rock, 1998; O'Shea & Fieo, 2014; Simons & Jensen, 2009).

The second type of task is a laboratory-based task where the main task and/or the CS are realistic, like the well-known gorilla video in which subjects are asked to count passes as a gorilla walks through the scene, or Neisser & Becklen's hand game vs. ball game video in which a woman carrying an umbrella walks through. These studies have shown ranges of inattentional insensitivity between 10% (Dalton & Fraenkel, 2012) for a study in which a male voice saying "I'm a gorilla" is heard in the midst of a simulated party scene, and 93% (Dattel et al., 2013) for a study in which participants building with blocks didn't notice increased heat on their hands (see also Becklen & Cervone, 1983; Clifasefi, Takarangi, & Bergman, 2006a; Graham & Burke, 2011; Neisser & Becklen, 1975; Seegmiller, Watson, & Strayer, 2011a; Simons & Chabris, 1999).

The final version is one in which participants are asked to complete a real-world task in which an unexpected event or object is embedded. For example, in one study participants were asked to follow a jogger through a park, counting as he taps his head with either hand. Participants generally then miss a rough fight being staged by confederates (about 50% inattentionally insensitive: Chabris, Weinberger, Fontaine, & Simons, 2011). In another study designed to test whether a father could have been inattentionally blind to his children asleep in the backseat of a vehicle when he was angry and intoxicated, 90% of participants were inattentionally insensitive to the presence of simulated children in the backseat of a car (Arndt, Wood, Delahunt, Krauss, & Wall, 2006).

Other examples of this type include a study with doctors screening for cancer in lung CT scans who miss a gorilla picture embedded in the lung 92% of the time (Drew, Võ, & Wolfe, 2013) and participants driving and completing a navigation task who miss a blatant no left turn sign (80% inattentionally insensitive: Kennedy & Bliss, 2013). Finally, multiple observational studies involve an unexpected event along a normal walkway and look for any signs of noticing by pedestrians (39%-100% inattentionally insensitive: Cornell, 1959; Hyman, Boss, Wise, McKenzie, & Caggiano, 2010).

Ranges of inattentional insensitivity across methods and modalities vary greatly. It is particularly important to point out that inattentional insensitivity is (or can be) present in any given situation or modality. However, without using comparable methods it is extremely difficult to determine whether there are systematic differences in inattentional insensitivity for visual vs. auditory stimuli and in intramodal vs. crossmodal paradigms. In addition to variation due to differences in methods and modalities, variance is also observed between individuals. It is this variation that may explain the vast differences found in previous work where researchers manipulate aspects of the task but fail to account for individual differences in their subjects.

1.6 Individual Differences in Inattentional Insensitivity

Individual differences have long been known to play a role in attention and cognitive control. However, the field of inattentional insensitivity constitutes a particular challenge when it comes to investigating the contribution of individual differences. Specifically, this is because inattentional insensitivity paradigms rely on one-shot, *unexpected* events (Bredemeier & Simons, 2012) meaning that sample sizes must be relatively large. Despite this limitation, researchers have bravely plowed ahead. However, findings have been relatively varied. This is likely due, in part, to the discrepancy in both experimental and analytical methods and a lack of variation in paradigms. Researchers have alternately studied working memory capacity, emotional salience of the critical signal, behavioral and societal correlates, and chronological age in relation to inattentional blindness.

1.7 Working Memory Capacity and Inattentional Insensitivity

An important individual difference characteristic that has been explored in previous research is Working Memory Capacity (WMC). Working memory is broadly defined as a temporary store in which information is kept before either being discarded or encoded into long term memory (Baddeley, 2012; Baddeley & Hitch, 1974). Working memory capacity, by extension is a representation of how much an individual can hold in this temporary store at one time. Working memory capacity can be measured in various ways, generally assessing the ability of a component referred to as the central executive, or the component of working memory which allocates attentional resources (Kane & Engle, 2002; Wickens & Hollands, 2000). Often, this is measured using span tasks, which may be simple (tasks in which subjects are required to repeat back correctly lists of items) or complex (where to be remembered lists are interleaved with demanding secondary tasks (see Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Complex span tasks used to assess working memory capacity or executive function include reading span, operational span and counting span (Conway et al., 2005) and can be used to assess visual or verbal working memory (Daneman & Carpenter, 1980). Researchers have alternatively hypothesized that inattentional insensitivity might be present due to either a lack of executive control, over-control in the form of overactive inhibition (Hannon & Richards, 2010), or an overall maxing out of WMC.

Findings have been varied with some research finding no differences in WMC and inattentional insensitivity (Bredemeier & Simons, 2012; Simons & Jensen, 2009) and others finding that low WMC individuals show higher rates of inattentional insensitivity, almost always in the form of inattentional blindness (Hannon & Richards, 2010; Seegmiller, Watson, & Strayer, 2011). Bredemeier & Simons (2012) found that WMC (as measured by performance on verbal and spatial 2-back tasks and then by the OSPAN task) was unrelated to noticing in a lab-based computerized inattentional blindness paradigm. However, they included all participants in their analysis, not just those who performed the primary task accurately. This is in contrast with, in particular, Seegmiller, Watson & Strayer (2011) who found that, for participants who performed the primary task (counting passes in a video) accurately, WMC (as measured by OSPAN) was indicative of noticing, where high WMC individuals were more likely to notice the critical stimulus than were low WMC individuals. Importantly, they found that, for participants with incorrect pass counts, there was no difference in noticing. This suggests that the allocation of attention to the primary task is essential for this paradigm and that studies that neglect to account for primary task performance may explain some of the variance in findings across tasks. These two studies show the major differences in methods including task type (computerized and abstract or realistic), WMC measurement (n-back/OSPAN treated as either continuous or dichotomized) and data inclusion (all participants or only those performing the primary task accurately).

Finally, and somewhat conflictingly, Hannon and Richards (2010) found that, overall, individuals with low WMC were more likely to be inattentionally blind than were high WMC individuals (similar to the findings of Seegmiller, Watson and Strayer). But, Hannon and Richards also point out that they found multiple individuals with very high WMC who were, nonetheless, inattentionally blind. Hannon and Richards indicate that this may indicate a dual-route model of inattentional blindness: that individuals who are inattentionally blind may either have exhausted their resources (in the case of low WMC individuals) or may have overactive inhibition (in the case of high WMC individuals). This may, therefore, explain the non-significant findings of previous WMC studies. Namely, that there is a possibility that dichotomized groups or linear predictions may not capture variance existing at the high WMC end of the spectrum. Furthermore, individual WMC differences have not, thus far, been examined in the context of inattentional deafness or crossmodal inattentional insensitivity. If inattentional blindness is a result of low WMC in the form of exhausted resources, it would be predicted that this difference should not be present in the case of crossmodal paradigms as, theoretically, resources should only be exhausted in the primary task modality and remain intact in the CS modality.

1.8 Behavioral and Societal Correlates of Inattentional Insensitivity

Further studies have indicated that those who are mildly intoxicated show markedly higher rates of inattentional blindness than those who are sober, and that that effect can be explained only by actual intoxication not by merely being told they (participants) have had alcohol (Clifasefi, Takarangi, & Bergman, 2006). Clifasefi et al. suggest that the breakdown in cognitive control associated with mild intoxication was responsible for the high rate of inattentional blindness in their sample. Their theory is supports the alcohol myopia theory set forth by Steele & Josephs (1990) which attributes lesser ability to process and extract meaning from cues, even if they are perceived at a physical level. These theories together may imply that individuals with poorer executive control would be expected to show higher levels of inattentional blindness.

Another recently investigated area of study has included whether race and social goals may moderate inattentional blindness levels. Brown-Iannuzzi, Hoffman, Payne, & Trawalter (2014) systematically manipulated interpersonal goals for a large set of Caucasian women and then showed them videos similar to Simons and Chabris' (1999) gorilla videos, though, in the case of this study, instead of a gorilla crossing the basketball scene, either a Caucasian or an African American man walked through the scene. Brown-Iannuzzi et al. found that when women had been given an interpersonal goal (to look for a friend, a romantic partner, a coworker or a neighbor, levels of inattentional blindness were almost always higher for the African American man than for the Caucasian man. Particularly, in the looking for a friend manipulation, the women were significantly more likely to be inattentionally blind to the African American man. Finally, in the control condition, where participants were given no interpersonal goals, women were somewhat more likely to notice the African American man than the Caucasian man. The authors took these findings to imply that a) the manipulation of interpersonal or task goals might serve to reduce levels of inattentional blindness when they bear relevance to the critical stimulus and that b) the somewhat lower inattentional blindness for the African American man in the control condition might indicate that the women subconsciously viewed him as more threatening, and therefore were more likely to be aware of his presence relative to the Caucasian (same race) man. Although the authors did not include male or African American participants in their sample, the idea that subconscious fear may motivate individuals to be less inattentionally insensitive to threatening stimuli may have merit and has been further explored in studies manipulating the emotional salience of the CS.

1.9 Emotional Salience and Inattentional Insensitivity

In an interesting study titled, "Spiders at a Cocktail Party", New and German (2014) systematically varied the critical stimulus presented in an inattentional blindness task similar to that used by Mack & Rock (1998). Participants were split into groups with critical stimuli varying on both ancestral (evolutionary) threat level and on more recent threat level. Stimuli included simple templates of spiders (high ancestral threat level) both curvilinear and rectilinear versions, hypodermic needles (more threatening to us now) and then of jumbled versions of the pictures to determine the effect of specific features versus identifiable pictures. New and German found that participants were much less likely to be inattentionally blind and more likely to be able to locate and identify templates of spiders than of other stimuli (needles or jumbled pictures). The authors, again, imply that this decrease in inattentional insensitivity is due to the stimulus being threatening, and therefore, despite its task irrelevance, drawing attention more readily.

1.10 Age Differences in Inattentional Insensitivity

Older adults experience age related declines in both physical perception and in cognitive processing abilities. Age-related declines include less efficient processing, including processing speed, working memory capacity, inhibitory function and long-term memory retrieval (Lustig, Hasher, & Zacks, 2007, Park & Reuter-Lorenz, 2009). Sensory declines are often considered to be inextricably linked with information processing declines in older adults (Murphy, Craik, Li, & Schneider, 2000; Slawinski & MacNeil, 2002). For example, Murphy et al (2000) found that the presence of noise adversely affects older adults' abilities to detect auditory signals and Slawinski & MacNeil found similar effects when using music to mask auditory warnings. These findings may indicate a greater deterioration of physical detection ability in the presence of noise or may indicate that older adults are more distracted by irrelevant secondary sounds or tasks (Lindenberger & Mayr, 2013). This hypothesis, of greater distractibility, has been posited to be related to age-related declines in working memory capacity (Park & Reuter-Lorenz, 2009). Specifically, Park & Reuter-Lorenz hypothesize that older adults encode stimuli inefficiently in working memory, processing and encoding not only relevant but also irrelevant signals, and may not "delete" irrelevant information from longer term working memory stores. They further explain that much of the variance found in age-related declines was mediated by working memory capacity and processing speed. Additionally, Baldwin & Ash, (2011) showed that age-related declines in sensory abilities can obfuscate results for tasks which are designed to measure WMC in older adults (such as the listening-span or L-span). Baldwin & Ash showed that older adults' L-span scores decreased with presentation level (from 65 dBA to 45 dBA) while younger adults' scores also decreased but not to the same extent. These findings, along with others (see Keidser, Ronnberg, Hygge, & Rudner, 2014; Lindenberger & Baltes, 1994; Valentijn et al., 2005) indicate that it is particularly important to equate stimuli based on sensory ability.

Of these well-documented changes, some may be of particular interest to inattentional insensitivity researchers. Few studies have specifically investigated the topic of inattentional insensitivity in older adults. However, one such exception includes the work of Graham & Burke (2011) who used an inattentional blindness paradigm with a sample of older adults in order to address two models of cognitive aging: attentional capacity and inhibitory deficit models. Cognitive models of attentional capacity assume that attention is finite and must be allocated or shared between tasks (see Kahneman, 1973). Graham & Burke point out that attentional capacity models have been used to explain age-related declines in cognitive performance by inferring that aging decreases attentional capacity, meaning that older adults should have less attention overall to allocate, particularly in the case of dual-task paradigms. This model would predict that older adults should show greater inattentional blindness than younger adults if the central task is equally difficult because older adults would have less spare attention to allocate to the unexpected stimulus. Conversely, Graham and Burke point to inhibition deficit models which postulate that aging decreases one's ability to inhibit processing of irrelevant information (see Zacks, 1989, Hasher & Zacks, 1988). This model would predict that older adults should show less inattentional blindness than younger adults because they would be unable to inhibit processing of irrelevant stimuli (theoretically to the detriment of the central task). Using the old standard video of basketball passes with a gorilla as the critical stimulus, they found that older adults noticed the gorilla significantly less than did younger adults. They interpreted this as support for attentional capacity models of cognitive aging. Specifically, they concluded that older adults simply lack the attentional resources to attend to the critical signal.

1.11 Physiological Measures in Inattentional Insensitivity Paradigms

Important in inattentional insensitivity paradigms is the idea that, although critical stimuli are not reported by subjects (explicit perception), they may have been *implicitly* perceived but not have ever reached conscious awareness. This theory is easily tested in the visual domain, using eye-tracking. For example, in a study with radiologists, Drew, Võ, & Wolfe (2013) found that 20 of 24 radiologists failed to report a picture of a gorilla embedded in CT lung-cancer slices but eye tracking data revealed that 12 of the 20 who did not report the gorilla, actually looked directly at its location (see **Error! Reference source not found.** for an example)

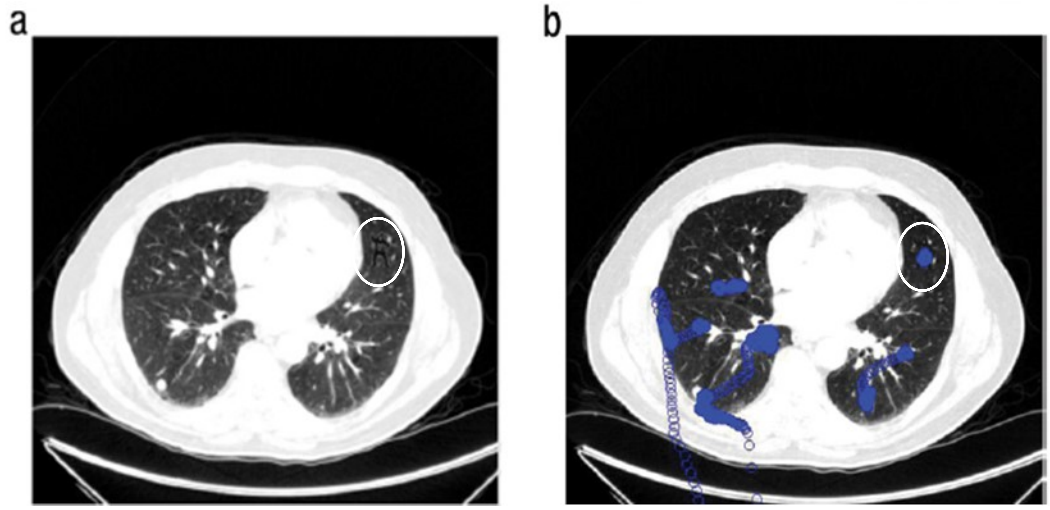

Figure 3. CT image containing the embedded gorilla (a, gorilla is in upper right quadrant and is circled in both pictures) and eye-position data for a radiologist who reported not noticing the gorilla (b) (Drew et al., 2013).

Additionally, an eye-tracking study by Memmert (2006), in which participants watched the well-known gorilla video, reported no significant differences in number of fixations, duration of fixation or average fixation on the gorilla between those who reported noticing and those who reported not noticing the gorilla.

1.12 Inattentional Insensitivity and Driving

Inattentional insensitivity has also been found to be present naturally in a variety of real-world tasks, in particular driving. Perceptual errors have been found to be highly related to roadway accidents (Galpin, Underwood, & Crundall, 2009; Nagayama, 1978; White & Caird, 2010). Looked but failed to see (LBFTS) accidents represent one of the commonly reported causes of collisions (Treat, 1980; White & Caird, 2010). LBFTS accidents occur when the driver is not visually or aurally impaired but is still unable to perceive a hazard. These crashes have been shown to be more likely when drivers are distracted, for example, by a conversation with an attractive passenger (White & Caird, 2010) or by a telephone conversation (Scholl, Noles, Pasheva, & Sussman, 2003; Strayer, Cooper, & Drews, 2004; Strayer, Drews, & Johnston, 2003). In order to improve driver behavior and decrease accidents, researchers and vehicle designers have focused on the implementation of advanced driver assistance systems (ADAS). In theory, these systems should act as a supplemental perceptual system and be able to alert a driver to the presence of a hazard, such that accidents due to perceptual errors, like LBFTS accidents are eliminated entirely. The first step towards effective ADAS design, is appropriate signal design.

1.13 Signal Design

With the recent proliferation of portable, handheld and wearable devices there has come an increasing need for well designed, unambiguous signals. This need is present throughout our daily lives: the need to distinguish between a call and an alarm from a vibrating phone, to distinguish between walk and don't walk in a crosswalk, to understand the meaning of a tornado alarm or to locate the direction of an ambulance siren. The pervasion of visual, auditory and tactile signals in our everyday life is rarely more obvious than in our vehicles. Automobiles can now not only tell you when you drift out of your lane, but can correct it. They can tell you the weather, the traffic conditions, your schedule, the status of your engine, your oil and your tires. They can tell you where to go, when to go and when to stop. They can do this visually, aurally and, in some cases, tactually. This plethora of information can come at a price though. When signals are unclear, ambiguous or poorly presented, they may cause distraction or confusion or be missed completely making the already complex task of driving even more difficult. Thus, it is increasingly important to develop good, appropriate and well mapped signals for in-vehicle applications.

This method starts at the very bottom- with the evaluation of perceptions of variable parameters. One of the most important perceptions is that of importance or urgency. Urgency relates to the hazardousness of the situation but is also implicit in physical signals. Visual signals that flash quickly or auditory or tactile alerts that have fast tempos have been shown to have higher levels of implicit urgency even without being matched to a scenario in context (Baldwin & Lewis, 2014; Baldwin et al., 2012; Hellier, Edworthy, & Dennis, 1993; Jang, 2007). Urgency scaling, or the development of continuous scales as units of measurement for urgency for various parameters is based on psychophysical scaling developed by Stevens in the 1950s (Hellier, Edworthy & Dennis, 1993). Psychophysical scaling involves the matching of continuous physical parameters such as auditory frequency or tempo, visual color or size, and tactile intensity or duration to other continuous parameters like brightness (crossmodal matching) or to numbers (magnitude estimation: Stevens, 1957). Developing urgency scales allows designers to more appropriately map alerts to situations, termed “urgency mapping” (Hellier & Edworthy, 1999). Urgency mapping ensures that highly salient, possibly annoying or abrasive alerts are only used in highly critical situations and that low saliency alerts are only used to signify information that is not time-critical and or is unimportant to safe vehicle operation.

Interestingly, despite the increasing implementation of appropriate and well-designed signals, warnings and alarms are still missed in highly critical situations leading to the exact perceptual errors they were designed to eliminate. Dehais and colleagues (2012, 2014) investigated the phenomenon whereby pilots miss highly critical, highly salient alarms during periods of particularly high workload. Dehais et al. found that, during an approach scenario with a high windshear (drastically increasing the difficulty of the landing maneuver), about 40% of pilots missed alarms indicating that their landing gear had malfunctioned. This leads to the question: why?

2 STUDIES ON INDIVIDUAL RESEARCH QUESTIONS

The following series of studies investigated perceptions of the urgency of alerts in multiple modalities, behavioral responses to highly urgent alerts in an in-vehicle context, and the effect of modality compatibility and individual differences on sensitivity to alerts during a perceptually demanding task. The intention of the detailed series of studies was to investigate the parameters of effective alerts, validate their effectiveness, and then to give some rationale for why alerts which should be effective, are sometimes still missed by operators.

2.1 Perceived Urgency Mapping across Modalities within a Driving Context: Study 1 Overview

The first study used psychophysical scaling in the form of numerical estimations of urgency and annoyance, for auditory, visual, and tactile signals to estimate comparative utility for various parameters of each modality.

2.1.1 Perceived Urgency Mapping across Modalities within a Driving Context

Baldwin, C. L., & **Lewis, B. A.** (2014). Perceived urgency mapping across modalities within a driving context. *Applied ergonomics*, *45*(5), 1270-1277.

2.1.2 Abstract

Hazard mapping is essential to effective driver vehicle interface (DVI) design. Determining which modality to use for situations of different criticality requires an understanding of the relative impact of signal parameters within each modality on perceptions of urgency and annoyance. Towards this goal we obtained psychometric functions for visual, auditory and tactile interpulse interval (IPI), visual color, signal word, and auditory fundamental frequency on perceptions of urgency, annoyance, and acceptability. Results indicate that manipulation of IPI in the tactile modality, relative to visual and auditory, has greater utility (greater impact on urgency than annoyance). Manipulations of color were generally rated as less annoying and more acceptable than auditory and tactile stimuli; but they were also rated as lower in urgency relative to other modality manipulations. Manipulation of auditory fundamental frequency resulted in high ratings of both urgency and annoyance. Results of the current investigation can be used to guide DVI design and evaluation.

Keywords: auditory warnings, tactile warnings, visual warnings, driver-vehicle interface, perceived urgency, annoyance, psychometric function, multimodal comparisons

2.1.3 Introduction

Modern automobiles, like many other advanced technological systems, utilize increasingly sophisticated displays capable of presenting information to the driver in a variety of ways. New components continue to be introduced into the driver vehicle interface (DVI) increasing both its potential usefulness and complexity. One of the many advances that have taken place recently is the use of vibrotactile signals in addition to the more common auditory and visual displays. Vibrotactile signals show promise for improving a driver's response to potential collision situations, particularly under distracted conditions (Fitch, Hankey, Kleiner, & Dingus, 2011) and when presented in combination with signals in other sensory modalities (Ho, Santangelo, & Spence, 2009; Lee, McGehee, Brown, & Marshall, 2006).

Determining the relative merit of providing information to drivers in one modality versus another is a challenging task. The choice of which modality to use will depend on many factors including the context in which the signal is likely to occur (e.g., daylight driving in dense traffic in relatively noisy conditions versus nighttime driving in quiet surroundings), driver characteristics (e.g., driver's age and sensory/cognitive capabilities), as well as the driver's state (e.g., alert versus fatigued) and habits (e.g., likely to be engaged in multiple tasks and distractions) and experience level (expert versus novice drivers). Also critical is the situation that the cue is designed to represent. Current and future driver DVIs will provide signals to drivers that are designed to represent a considerable range of situations that vary in criticality and importance. For example, current systems such as SYNC for MyFord Touch®[100] or MyLincoln Touch®[122] support hands free interactions with incoming calls and texts and navigation assistance while also interacting with driver safety systems such as blind spot indicators and collision avoidance technologies. It is critical to effective design that signals are appropriately mapped to the situations they are designed to represent (Dingus, Jahns, Horowitz, & Knipling, 1998; Edworthy, 1998; Edworthy, Loxley, & Dennis, 1991; Edworthy & Stanton, 1995).

Perceived Urgency Mapping

Mapping the perceived urgency of a signal to the hazard level which it is designed to represent has been recognized as an important aspect of warning design since at least as far back as the 1980's (Chapanis, 1994; Edworthy et al., 1991; Hollander & Wogalter, 2000; Patterson, 1982, 1990; Wogalter & Silver, 1990, 1995). When signals are too prevalent, intense, abrasive, startling, or simply too numerous they cause annoyance and distraction (Baldwin, 2011; Edworthy et al., 1991; Marshall, Lee, & Austria, 2007; Wiese & Lee, 2004), have little or no performance benefit (Baldwin & May, 2011), reduce trust in the system (Lees & Lee, 2007) and can even lead to impaired reactions to subsequent critical events (Fagerlonn, 2011). As the number of displays and alerts in the DVI proliferate it will be increasingly important to ensure that the signals, alerts and warnings presented convey appropriate levels of urgency.

The relationship between changes in several key physical stimulus parameters and perceived urgency in visual and auditory modalities has been documented. For example, as the fundamental frequency of a sound increases, and/or as the time interval between pulses of sound decreases, it is perceived as increasingly urgent (Edworthy et al., 1991; Hellier & Edworthy, 1999; Hellier, Edworthy, & Dennis, 1993). Likewise, as the wavelength of visible light increases (hue or perceived color changing from green to yellow to red) it is perceived as more urgent (Chapanis, 1994; Wogalter, Conzola, & Smith-Jackson, 2002; Wogalter, Kalsher, Frederick, Magurno, & Brewster, 1998). In general, there is a direct relationship between perceived urgency and annoyance, such that as a signal becomes more urgent it is also perceived as more annoying (Baldwin, 2011; Marshall et al., 2007). However, the context in which the signal is presented influences this relationship (Wiese & Lee, 2004). More urgent signals are perceived as less annoying in conjunction with situations where the high urgency seems appropriate (collision warnings) relative to situations where it is less appropriate to receive a very urgent signal (e.g., navigation command or email alert) (Marshall et al., 2007). Further research is needed to elucidate the impact of different types of context on the relationship between perceived urgency and annoyance and the potential impact that signal modality may play in this relationship. Choosing an effective modality and parameter level is particularly important for time critical situations represented by collision warnings.

Collision Avoidance Systems

Several research investigations have compared the time drivers take to respond to collision warnings presented in one modality versus another (Kramer, Cassavaugh, Horrey, Becic, & Mayhugh, 2007; Mohebbi, Gray, & Tan, 2009; Scott & Gray, 2008). For example, Scott and Gray (2008) compared brake response times to visual, auditory, and tactile collision warnings and concluded that tactile warnings resulted in the fastest responses. Mohebbi, et al. (2009) compared auditory versus tactile warnings when drivers were engaged in simple versus complex simulated cell phone conversations. Participants in their study exhibited faster response times to tactile warnings relative to auditory warnings while engaged in both simple and complex conversations. However, despite careful consideration of the type of auditory and tactile signals to present (gleaned from existing guidelines and the available literature), it remains possible that in both of these investigations the signals presented in the different modalities may not have been equally salient to drivers. That is, driver response times may have differed significantly from those observed had they used different types of auditory or tactile signals (e.g., different intensity, frequency, or temporal pulse patterns).

In fact, drivers may fail to even notice some visual warnings (Curry, Blommer, Greenberg, & Tijerina, 2009). This is of practical significance since an undetected alert is of little use. However, it is also possible that visual signals that are perceived as more urgent or that are more salient are more likely to capture attention and subsequently be more effective. For example, a flashing red alert that is perceived as highly urgent may be more effective than a low frequency, low intensity, long burst of sound. However, without first equating the two signals for perceived urgency it would be misleading to suggest that modality alone was driving differences in signal effectiveness. The primary rationale for the current study was to compare stimuli across visual, auditory, and tactile modalities for perceived urgency in order to facilitate future examinations of the effectiveness of modality across equivalent urgency levels.

A wide variety of different auditory, visual, and tactile signals have been compared. Visual signals frequently consist of an array of light emitting diodes (LEDs) in various colors (e.g., red, amber, yellow or green) that may or may not flash and may be located in a variety of head up and head down positions (Kramer et al., 2007; Neale, Perez, Lee, & Doerzaph, 2007; Scott & Gray, 2008) modeled after those examined in the Crash Avoidance Metric Partnership (CAMP) program (Kiefer et al., 1999). The CAMP program was designed to provide guidance on collision alert timing and modality requirements. Based in large part on that research, many subsequent investigations of auditory signals have examined various non-speech tones. In the CAMP project, Keifer et al. (1999) compared the crash warning capabilities of a tone with a peak at 2500 Hz and the spoken signal word, such as “Warning” repeated. Both the speech and non-speech auditory signals were set to play from the car speakers at 67.4 dBA. Kiefer et al. (1999) concluded that the non-speech tone had superior crash warning alert capabilities relative to the speech warning. This result corroborated Tan and Lerner’s (1995) multiattribute evaluation findings that the auditory sounds most likely to be effective as primary collision avoidance warnings were also nonverbal sounds. Many subsequent DVI researchers have tended to avoid speech warnings and concentrate instead on nonverbal tones. For example, Wiese and Lee (2004) compared two nonverbal sounds thought to convey different levels of urgency and Scott and Gray utilized a 2000 Hz tone.

It is of note that in the Kiefer et al. investigation, speech warnings were initially rated as the most favorable on key attributes, such as noticeability, and urgency. Despite this only one speech warning was examined and it was always presented in combination with a head down visual display. On average brake responses were slower to the speech warning relative to the non-speech warning. It is possible that the lower fundamental frequency of the speech or its relative ability to penetrate through the ambient background noise resulted in a fundamental difference in detectability and perceived urgency that could have significantly impacted the results. Various subsequent studies have found that acoustic factors interact with semantic factors (e.g., signal word) (see Baldwin & May, 2010 and Edworthy, Hellier, Walters, Clift-Mathews, & Crowther, 2003) and that this interaction can impact both perceived urgency and collision avoidance response (Baldwin, 2011).

Haptic or Tactile alerts vary, but in the CAMP report they consisted of a "vehicle jerk" that simulated the feeling of a brake pulse (Kiefer et al., 1999). Other researchers have examined vibrotactile signals presented in various places - the seat pan (Fitch et al., 2011) or a waist belt (Ho, Reed, & Spence, 2007; Ho et al., 2009; Mohebbi et al., 2009) at a variety of temporal rates. For example, Fitch et al., (2011) used an interpulse interval (IPI) of 50 ms in a collision avoidance signal context; Van Erp and Van Veen (2004) examined IPI rates ranging from 270 ms to 10 ms. Research for DVIs will need to determine the most effective parameters within each modality or combination of modalities for these imminent crash warnings while also examining efficient methods of cueing the driver's attention appropriately to less critical situations. Equating signals for urgency across differing modalities and modes will be essential to appropriate hazard mapping for both critical and noncritical alerts.

Noncritical Alerts

Not all signals presented to drivers should connote high urgency. For example, drivers may be alerted to incoming phone messages and emails, receive information regarding future and near turn route guidance, as well as weather, traffic, and road conditions. Future DVIs will have the capability of providing even more information to the driver making it critical to appropriately match the urgency conveyed with the importance and time criticality of the situation it represents. For example, in one investigation of driver acceptance of simulated distraction mitigation alerts both middle-aged and older drivers reported significantly higher acceptance of these non-time critical alerts when they were presented in a visual rather than auditory modality (Donmez, Boyle, Lee, & McGehee, 2006). The diversity of situations varying in time criticality requires careful attention to methods of effectively alerting drivers while minimizing annoyance. At present, however there is little research regarding how perceptions of urgency, annoyance and acceptability differ across modalities. In one notable exception Baldwin et al. (2012) developed psychometric scales of perceived urgency and annoyance for specific parameters in the visual, auditory, and tactile modalities. They found that a range of perceived urgency could be obtained in each modality. Manipulations of interpulse interval (IPI) in auditory and tactile modalities were particularly effective at altering perceptions of urgency. In particular, tactile IPI tended to exhibit the greatest range of urgency ratings while changes across the physical parameter had far less impact on ratings of annoyance relative to changes in IPI for auditory tones. Using the terminology of Tan and Lerner (1995) this parameter can be said to have good “utility.” Establishing a range of signal parameters that alert drivers while conveying relatively low levels of urgency is important for providing non-time critical events. However, previous research indicates that context can have a significant impact on ratings of urgency and annoyance (Marshall et al., 2007).

Context

Much of the early work examining the impact of different signal parameters on perceptions of urgency and annoyance was conducted with little or no contextual information and therefore could potentially be generalized to a variety of operational environments (Edworthy, 1994; Edworthy et al., 1991; Edworthy & Stanton, 1995; Wogalter & Silver, 1995). Asking participants to rate signals within the driving context (Baldwin, 2011) leads to similar trends as more generic settings. However, different relationships between ratings of urgency and annoyance have been found when examining signal parameters within specific contexts. For example, Marshall et al. (2007) examined ratings of urgency an annoyance for changes in auditory signal parameters in three different DVI contexts – collision avoidance, navigation and e-mail alerts. Annoyance ratings for several of the signal parameters investigated were affected by the context participants were asked to use. Specifically, for instance a sound with a longer IPI was rated as less urgent and less annoying than one with a shorter IPI. But, ratings of annoyance were affected less by the IPI manipulation in the e-mail context relative to the collision avoidance context. Baldwin et al.’s (2012) investigation included a low level of context. Participants were asked to provide magnitude estimations for a context within the vehicle while simply looking at a computer-generated image of a dashboard. No information was provided as to what the signal was supposed to represent (i.e., collision warning or e-mail alert). It is possible that drivers would perceive the signals differently if they encountered them while engaged in actual or simulated driving. We address this issue in the current investigation.

2.1.4 Current Investigation

In order to determine the relative merits of each modality, per se, for signaling events of different hazard levels it is critical that researchers implement a means of equating the signals being examined for their perceived urgency. Establishing and validating scales of perceived urgency across modalities was the primary goal of the current investigation. A second goal of the current study was to compare relative signal utility. In other words, we sought to examine changes in signal parameter manipulations on ratings of annoyance relative to changes in perceived urgency across visual, auditory and tactile modalities. Drawing from the existing literature, we hypothesized that perceived urgency would increase as auditory frequency increased and IPI decreased in both auditory and tactile modalities (Baldwin, 2011; Baldwin et al., 2012; Hellier et al., 1993; Marshall et al., 2007) and that perceived urgency would be higher for the signal words “Danger” relative to “Warning” and “Notice” and as these words were presented in red relative to yellow and green backgrounds (Chapanis, 1994; Wogalter & Silver, 1995). Further, we hypothesized that changes in IPI would result in greater signal utility (have a greater impact on urgency than annoyance) relative to changes in fundamental frequency or color. There is a relative lack of research regarding the impact of changes in tactile parameters on perceptions of urgency and annoyance (but see Baldwin et al., 2012). However, we hypothesized that the tactile modality would yield greater signal utility relative to auditory and visual modalities based on previous work (see Lewis & Baldwin, 2012).

2.1.5 Materials and Methods

Participants

Participants were 20 (7 male, average age = 23.9 years, 16 right-handed) graduate and undergraduate George Mason University students. Participants were recruited via the psychology research pool or through word of mouth. Participants participated either on a volunteer basis or, if desired, were compensated via the University's research participation credit system. All participants were given a university human subjects review board approved written informed consent document and acknowledged their voluntary participation by signature. All participants self-reported normal or corrected to normal vision and hearing.

Apparatus and Materials

Participants were seated in front a desktop driving simulator using a Logitech racing wheel and pedal set integrated with a medium fidelity simulation generated by Realtime Technologies, Inc software. The experimental stimuli were played on a Dell Duo laptop located directly behind the steering wheel (where the dashboard would normally be located in a vehicle), underneath the simulator screen as shown in Figure 4.

Figure 4. Experimental setup: desktop driving simulator and presentation laptop location.

Throughout the course of the experiment, the background image on the presentation laptop was a generic car dashboard. Stimuli were presented using a custom-built MATLAB program in conjunction with Psych Toolbox. Auditory stimuli were played through the speakers and set to be presented at an average of 20 dB (measured using a Brüel & Kjær Sound Level Meter- Model 2250) above the ambient road and engine noise coming from the simulator when the vehicle was at moving at 35 mph. Stimuli were set to be 20 dB above ambient noise based on recommendations made by Patterson & Mayfield (1990). Tactile stimuli were played via sound files through the computer using a single C2© (Engineering Acoustics, Inc.) tactor and a RadioShack amplifier modified to act as a microcontroller. Visual stimuli were presented embedded in a simulated dashboard as illustrated in Figure 5.

Stimuli

Auditory

Auditory stimuli were created in Adobe Audition CS 5.5. Auditory stimuli consisted of signals with seven different inter pulse interval (IPIs) and seven complex sounds varying in fundamental frequency (F0). The full list of the stimuli presented is provided in Table 2.

Table 2. Complete list of auditory, visual and tactile stimuli examined.

<table>
<tr><td colspan="8">Auditory</td></tr>
<tr><td>IPI (ms)</td><td>475</td><td>302</td><td>238</td><td>118</td><td>60</td><td>50</td><td>9</td></tr>
<tr><td>Frequency (Hz)</td><td>210</td><td>250</td><td>260</td><td>320</td><td>440</td><td>500</td><td>680</td></tr>
<tr><td colspan="8">Visual</td></tr>
<tr><td>IPI (ms)</td><td>475</td><td>302</td><td>238</td><td>118</td><td>60</td><td>50</td><td>9</td></tr>
<tr><td>Color (nm)</td><td colspan="2">510
(Green)</td><td colspan="2">580
(Yellow)</td><td colspan="2">608
(Orange)</td><td>645
(Red)</td></tr>
<tr><td>Word</td><td colspan="2">Notice</td><td colspan="2">Brake</td><td colspan="2">Warning</td><td>Danger</td></tr>
<tr><td colspan="8">Tactile</td></tr>
<tr><td>IPI (ms)</td><td>475</td><td>302</td><td>238</td><td>118</td><td>60</td><td>50</td><td>9</td></tr>
</table>

All signals had a total duration of approximately 2500 ms with 200 ms pulse durations (with 20 ms onset and offset times) and were modeled after those examined by Hellier et al. (1993). Total stimulus lengths varied slightly when IPI was manipulated because the signal ended if the last pulse ended before 2500 ms without sufficient time to include another IPI and pulse. IPI times ranged from 475 ms (being the slowest IPI stimuli) to 9 ms (being the fastest). When IPI was being manipulated, the F0 of the signal was 300 Hz with 15 harmonic components, stimuli did not have any IPI time but included 20 ms onset and offset times so that signals would not be perceived as a constant pulse. Frequency stimuli consisted of similar sounds with F0s ranging from 210 Hz to 680 Hz with 15 harmonic components.

Visual

Visual stimuli consisted of seven flashing stimuli (referred to here as IPI stimuli; the visual equivalent of IPI stimuli in the auditory or tactile modalities), four color backgrounds and four signal words. Visual IPI stimuli were manipulated identically to auditory IPI stimuli. When IPI was being manipulated the stimulus flashed at the rate indicated in Table 2 and the signal word “Warning” was presented in black ink on a background of yellow as illustrated in Figure 5.

Figure 5. Visual stimulus example.

When IPI was not being directly manipulated the stimuli did not flash but were presented for a duration of 2500 ms. Color stimuli were created on a Dell Laptop and transformed to wavelengths (reported in nanometers). Colors ranged from 510 nm (green) to 645 nm (red). When color was manipulated, the signal word “Warning” was presented. The signal word was always presented in black unless the background color was red, then white was used. Previous research indicates that white is more visible than black on the red background color (Laughery, 2006). As listed in Table 2 the four signal words examined were “Notice”, “Warning”, “Danger”, and “Brake”. When signal word was being manipulated, words were presented in black ink against the background color yellow.

Tactile

Tactile signals were created via Adobe Audition CS 5.5 and consisted of the same basic pulse level (300 Hz) as auditory stimuli but without any harmonic components. Tactile signals varied in IPI in a manner identical to the auditory stimuli. All IPI levels are listed in Table 2.

Procedure

Participants were seated in a chair facing the driving simulator in a sound attenuated laboratory room. After providing written informed consent, they completed a short demographic survey and then were directed to the simulator. Participants were allowed to adjust their seating until they felt comfortable with the position of both the steering wheel and the pedals.

The researcher then attached the tactor to the top of the participant's left forearm similar to the methods used by Ferris & Sarter (2008). The tactor was placed over a piece of store brand plastic wrap (in order to keep the tactor from contact with uncovered skin), approximately one inch above their wrist joint and secured in place using an athletic sweat band. Participants were informed that they would receive stimuli in each of the three modalities in separate blocks while they drove on a long continuously looping highway. Stimuli were presented in randomized order within each block and block order was counterbalanced between participants. Participants were instructed to maintain a target speed of 35 mph and to keep their simulated position within the lane to the best of their ability. Participants were then instructed to rate the stimuli they saw, heard or felt based on their subjective opinions of the urgency, annoyance and acceptability.

Acceptability was defined for them as their willingness to own or operate a vehicle with a similar alert. No further contextual information was provided regarding the nature of the stimuli or what they were designed to represent. Participants were instructed to continue driving at all times.

Participants were then given a practice session in order to familiarize themselves with driving using the simulator while receiving and rating stimuli. Participants drove until they were capable of maintaining their speed within approximately 5 mph around the target of 35 mph and maintain their simulated position within the lane. They then practiced driving while receiving and rating stimuli in all three modalities. Once participants had demonstrated their capability of maintaining driving performance while rating stimuli in the practice session, the experimental session was initiated. Practice stimuli consisted of a subset of 2 stimuli per modality from the middle values for each parameter. Participants were given the option to take a break between blocks if they wanted.

Design and Theoretical Calculation

Stimuli were chosen based primarily on previous work and physical parameter restrictions. All IPI and frequency values were modeled after those used by Hellier, Edworthy and Dennis (1993). All other stimulus values were based on previous work in our labs (see Baldwin et al., 2012). Both modality and parameter were manipulated within subjects. All subjects received all visual, tactile and auditory stimuli and rated them based on their perceptions of urgency, annoyance and acceptability on a scale of 0 (low urgency, low annoyance and unacceptable) to 100 (highly urgent, highly annoying and acceptable) similar to the procedure used by Marshall et al. (2007). This data was then examined to determine whether it met the assumptions of linearity according to the criterion established by S.S. Stevens (1957). We chose not to use his specific exponent-based power law function, because of range issues addressed by Teghtsoonian (1973). Teghtsoonian has demonstrated that differences in range size impact psychometric exponents. Specifically, use of a larger range will result in a smaller exponent and vice versa when estimation stimuli are limited (in this case between 0 and 100). While we did not use a power law exponent we did compare effectiveness on the basis of the slope of the best fit line for each parameter, directly related to Stevens' Power Law exponents (calculated as the slope of the best fit line through the log-log plots of stimuli). We chose not to log transform our variables since log transformed values weight lower values more heavily (e.g., our IPI stimuli would be weighted more heavily than our color stimuli). Instead we compared variables based on their relative changes in percentages of their individual, dynamic, ranges (such that the lowest value in each range was its 0% point and the highest was its 100% point) in order to resolve these issues. In this way, the raw rating values can be retained. There is a large body of evidence indicating that judgments of magnitude on one dimension can be reliably fit by a power law function of stimulus intensities without the necessity to use a logarithmic transformation (Teghtsoonian, Teghtsoonian, & DeCarlo, 2008; Teghtsoonian, 1973; Teghtsoonian, 2012).

2.1.6 Results

Perceived Urgency

Mean group ratings of perceived urgency for each parameter in each modality are illustrated in Figure 6 and Table 3.

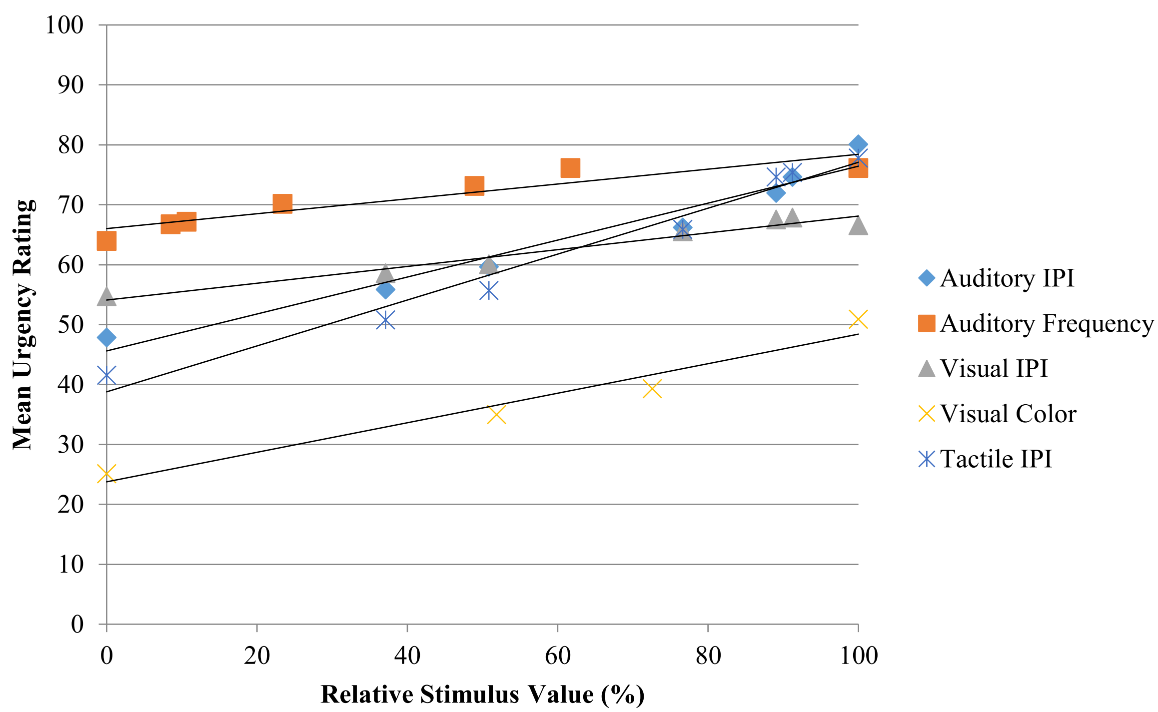

Figure 6. Average urgency rating of physical stimuli parameters, best fit lines are included for all parameters.

Increases in auditory fundamental frequency and decreases in interpulse interval (IPI) in each modality resulted in increased magnitude estimations of perceived urgency as illustrated by the positive slopes of the best fit lines.

Table 3. Slopes obtained from average urgency ratings of physical stimulus parameters (word values are excluded from all slope reports as their slopes cannot be calculated by these methods).

Urgency			
	Slope	Adjusted R^2	Significance
Auditory IPI	0.31	0.95	0.00
Auditory Frequency	0.12	0.84	0.00
Visual IFI	0.14	0.95	0.00
Visual Color	0.25	0.93	0.00
Tactile IPI	0.38	0.97	0.00

It is also important to note that the adjusted R^2 values for the best fit lines were always quite high (between .84 and .97) meaning that the data can be viewed as having a linear relationship between perceived urgency and relative stimulus magnitude. The degree of fit between obtained and predicted lines were all significant at less than a 0.001 level.

Annoyance

Figure 7 and Table 4 illustrate the mean annoyance ratings associated with each of the manipulated parameters.

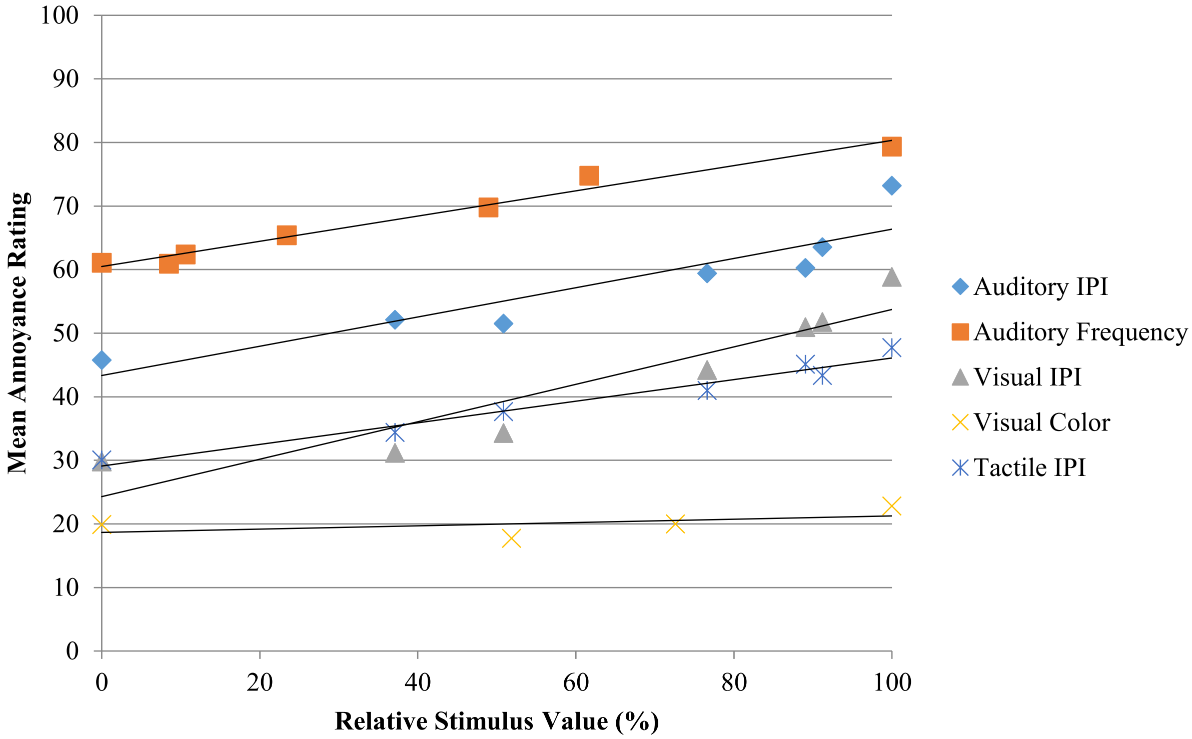

Figure 7. Average annoyance rating of physical stimuli parameters, best fit lines are included for all parameters.

Similar to ratings of perceived urgency, ratings for annoyance were found to have generally positive slopes, indicating that increases in auditory fundamental frequency and decreases in IPI in each modality resulted in increased magnitude estimations of annoyance. These relationships as illustrated in Figure 7 and resulted in positive slopes for the best fit lines.

Table 4. Slopes obtained from average annoyance ratings by relative physical stimulus parameters.

Annoyance			
	Slope	Adjusted R^2	Significance
Auditory IPI	0.23	0.80	0.00
Auditory Frequency	0.20	0.97	0.00
Visual IFI	0.29	0.84	0.00
Visual Color	0.03	-0.09	0.47
Tactile IPI	0.17	0.96	0.00

The relationship between observed and predicted values resulted in adjusted R^2 ratings between .80 and .87 and all were significant at or below the .001 significance level indicating that as the auditory fundamental frequency increased and IPI decreased for signals in each modality, perceived annoyance increased as well. It can also be noted that ratings for visual color resulted in an insignificant line that trended to be negatively sloped meaning that as urgency increased by visual color, annoyance either did not increase or in some cases actually decreased.

Acceptability

In general, acceptability ratings produced negative or insignificant slopes. However, for visual color the ratings of acceptability showed a significantly positive slope, meaning that even though the value of the physical parameter increased and the urgency ratings for that parameter increased, subjects still continued to find the stimuli more and more acceptable.

Urgency versus Annoyance

Figure 8 illustrates the relationships between urgency and annoyance ratings for all parameters.

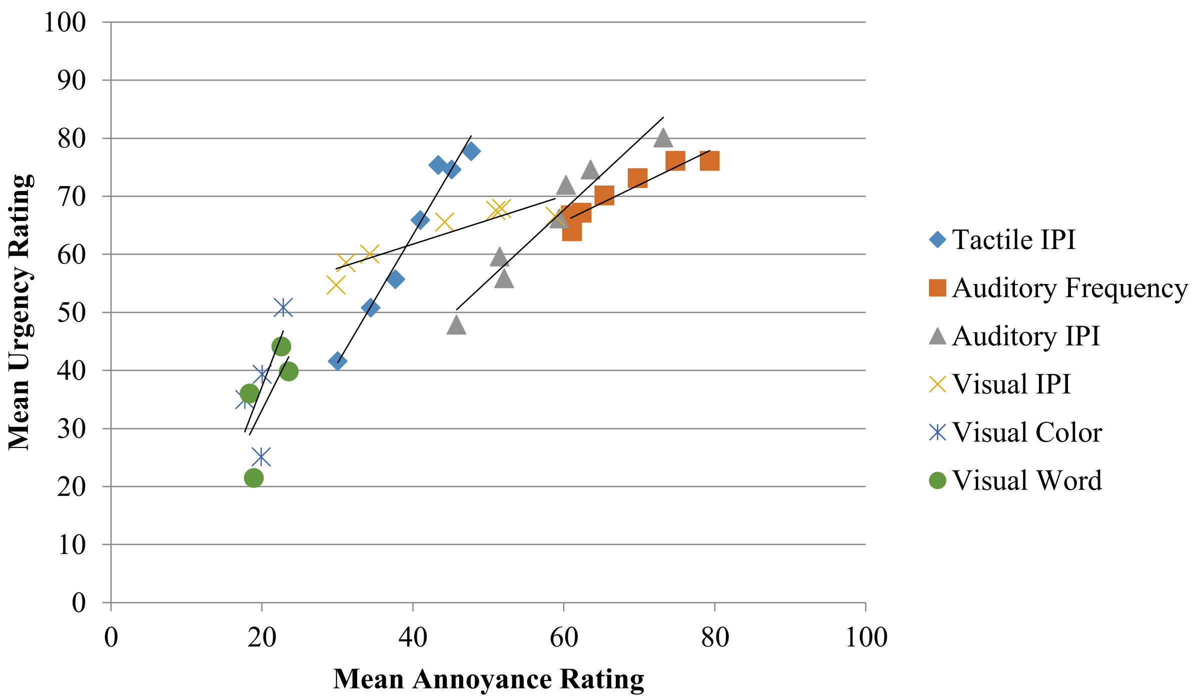

Figure 8. Mean urgency ratings plotted against mean annoyance ratings.

Table 5 reports the slopes of the best fit lines between all urgency versus annoyance data points. The greater the slope (or the more vertical the line) the greater the signal utility –

more urgency can be increased without increasing annoyance (as in the case of signal word, visual color, tactile IPI and auditory IPI).

Table 5. Slopes calculated from the best fit line between mean urgency and mean annoyance ratings.

Modality	Slope
Tactile IPI	2.21
Auditory Frequency	0.63
Auditory IPI	1.21
Visual IPI	0.42
Visual Word	2.59
Visual Color	3.41

Lower slopes or flatter lines in this case indicate lower signal utility indicating that even small changes in urgency result in larger changes in annoyance (as in the case of visual IPI and auditory F0). The length of the line provides a graphical depiction of the range of urgency levels that can be achieved by manipulating each signal parameter. Longer lines indicate a greater range of urgency ratings. Additionally, this graph allows comparison of the ranges of urgency between modalities as they relate to ranges of annoyance. For example, the urgency range covered by tactile IPI (about 42 to 78) covers nearly the same urgency range as auditory IPI (about 48 to 80) but results in much lower rates of annoyance (tactile IPI covers between 30 and 48 while auditory IPI covers from 45 to 73).

2.1.7 Discussion

A major goal of this investigation was to examine the relative impact of changes in specific signal parameters in visual, auditory, and tactile modalities on perceptions of urgency and annoyance. Specifically, we sought to determine the range of perceived urgency that could be obtained from signals in each modality and their relative utility in terms of the relationship of changes in urgency to changes in annoyance (Tan & Lerner, 1995). Our results for the visual and auditory modality are in line with existing literature (Baldwin, 2011; Baldwin et al. 2012; Chapanis, 1994; Hellier et al., 1993; Wogalter & Silver, 1995). Our results extend previous work to the tactile modality and facilitate comparison across modalities. As hypothesized, both perceived urgency and annoyance increased as frequency in the auditory modality and as IPI decreased in visual, auditory and tactile modalities. Additionally, perceived urgency was higher for the signal words "Danger" relative to "Warning" and "Notice" and for words presented in the background color red relative to yellow and green. More importantly, the current results provide some of the first ever comparisons of the changes in perceptions of urgency across each of the modalities as a function of changes to the signal parameters and the relative utility of parameter manipulations in terms of an urgency versus annoyance tradeoff.

As predicted, changing the temporal characteristics of signals in the tactile modality (IPI) yielded the largest range of perceived urgency ratings while having less impact on annoyance ratings, relative to the same manipulation in the auditory modality. This observation supports the conclusion that tactile IPI has a greater signal utility (has a greater impact on urgency than annoyance) relative to IPI changes in visual and auditory modalities and relative to auditory fundamental frequency and color. These results are in line with our previous work (Baldwin et al., 2012), but extend them to a driving simulation context which can argued to have greater relevance for the for the current goal of establishing an empirical basis for DVI design guidelines.

There are several limitations worth noting in the current investigation. First, only one parameter was manipulated in the tactile modality and these signals were presented at only one location- on the wrist. However, unpublished research from our lab indicates highly reliable urgency ratings across different body locations for tactile stimuli. We therefore predict that tactile pulses of equal intensity presented in other location would yield highly similar results. Results indicated that the tactile modality has greater utility than the auditory and visual parameters examined in the current investigation. However, since only IPI was manipulated here, it remains to be seen if other tactile signal parameters or tactile signals presented at different locations on the body will have similar effectiveness. Secondly, the current investigation examined only unimodal signals. There is a strong trend in DVI design to include bimodal signals. Further research regarding the perceived urgency and annoyance of signals presented in bimodal combinations is warranted. Additionally, though substantial ranges of urgency were obtained with the signal parameter manipulations examined here the results are limited to subjective ratings. Behavioral responses may differ from perceptions of urgency and this possibility must be examined in future DVI research. In the current investigation, we did not formally screen participants for audiometric, visual, or tactile acuity but rather relied on self-reported normal abilities. As in the general population, acuity varies across individuals and modalities. It is possible that some participants may have had undocumented acuity deficits or sensitivities. However, the within-subject nature of the experiment and the multiple presentations of each stimulus should have minimized the impact of potential individual differences. Further, although we only used college students, previous research utilizing psychophysical measurement has found that inexperienced college students are capable of performing as well as experts on these types of perception tasks and that these types of psychophysical effects can be found across populations. (Stevens, 1971; Teghtsoonian, 1973). Finally, though a simulated driving context was utilized in the current investigation it remains to be seen if providing a more specific context will impact both perceptions of and responses to signals in each of the modalities.

Despite these limitations results of the current study provide empirical evidence that can be used to guide DVI design. As previously discussed, inappropriately mapped (e.g., overly urgent or highly annoying and frequent) alerts and warnings can reduce trust, lead to little or no performance benefit and can even impair subsequent responses (Baldwin & May, 2011; Fagerlonn, 2011; Lees & Lee, 2007; Marshall et al., 2007). The current results provide key information on the utility of several parameters within multiple modalities. This information can be used to inform future DVI design and evaluation. For example, the relatively low ratings of both perceived urgency and annoyance for visual signals varying in color indicate that manipulations of color can be used to indicate non-time critical changes in system states, such as infotainment and personal communication notices. On its own, however, color would not be perceived as sufficiently urgent to signal more critical changes such as lane departures or collision events. Conversely, both tactile and auditory signals resulted in a sufficient range of urgency levels to be used to signal relatively critical information. However, given the higher ratings of annoyance for the auditory signals, the tactile modality would be preferred for signals with more frequent false alarms. Though conclusive guidelines require examination in higher fidelity driving contexts, the current results are expected to facilitate this effort.

2.1.8 Conclusion

Well-designed alerts and warnings can ensure that drivers receive the information they need to facilitate safe and efficient transportation. Signal modality can be an important method of assisting drivers with discriminating between the urgency of different situations. Determining which modality to use for a given situation requires empirical examination. But for that examination to be effective it is essential that issues of modality are not undermined by differences in signal parameters that affect urgency independent of modality. Results of the current investigation provide a means of ensuring that comparable levels or urgency can be achieved across different modalities to facilitate comparison of response patterns. Ratings obtained here indicate that manipulation of tactile IPI has greater utility than manipulations of auditory or visual stimulus parameters. These results can be used to provide guidance for effective DVI designs.

2.1.9 References

Baldwin, C. L. (2011). Verbal collision avoidance messages during simulated driving: Perceived urgency, alerting effectiveness and annoyance. *Ergonomics, 54*(4), 328-337. doi: 10.1080/00140139.2011.558634

Baldwin, C. L., Eisert, J. L., Garcia, A., Lewis, B., Pratt, S. M., & Gonzalez, C. (2012). Multimodal urgency coding: Auditory, visual, and tactile parmeters and their impact on perceived urgency. *Work, 41*, 3586-3591. doi: 10.3233/WOR-2012-0669-3586

Baldwin, C. L., & May, J. F. (2011). Loudness interacts with semantics in auditory warnings to impact rear-end collisions. *Transportation Research Part F-Traffic Psychology and Behaviour, 14*(1), 36-42. doi: 10.1016/j.trf.2010.09.004

Chapanis, A. (1994). Hazards associated with three signal words and four colours on warning signs. *Ergonomics, 37*(2), 265-275. doi: 10.1080/00140139408963644

Curry, R., Blommer, M., Greenberg, J., & Tijerina, L. (2009). Immediate Recall of Driver Warnings in Forward Collision Warning Scenarios. *Transportation Research Record*(2138), 28-33. doi: 10.3141/2138-05

Dingus, T. A., Jahns, S. K., Horowitz, A. D., & Knipling, R. (1998). Human factors design issues for crash avoidance systems. *Barfield, Woodrow*.

Donmez, B., Boyle, L. N., Lee, J. D., & McGehee, D. V. (2006). Drivers' attitudes toward imperfect distraction mitigation strategies. *Transportation Research Part*

F-Traffic Psychology and Behaviour, 9(6), 387-398. doi: 10.1016/j.trf.2006.02.001

Edworthy, J. (1994). The design and implementation of non-verbal auditory warnings. *Applied Ergonomics, 25*(4), 202-210. doi: 10.1016/0003-6870(94)90001-9

Edworthy, J. (1998). Warnings and hazards: An integrative approach to warnings research. *International Journal of Cognitive Ergonomics, 2*(1-2), 3-18.

Edworthy, J., Hellier, E., Walters, K., Clift-Mathews, W., & Crowther, M. (2003). Acoustic, semantic and phonetic influences in spoken warning signal words. *Applied Cognitive Psychology*, *17*(8), 915–933.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors, 33*(2), 205-231.

Edworthy, J., & Stanton, N. (1995). A user-centred approach to the design and evaluation of auditory warning signals: I. Methodology. *Ergonomics, 38*(11), 2262-2280.

Fagerlonn, J. (2011). Urgent alarms in trucks: effects on annoyance and subsequent driving performance. *Intelligent Transport Systems, 5*(4), 252-258. doi: 10.1049/iet-its.2010.0165

Ferris, T. K., & Sarter, N. B. (2008). Cross-Modal Links Among Vision, Audition, and Touch in Complex Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(1), 17–26. doi:10.1518/001872008X250566

Fitch, G. M., Hankey, J. M., Kleiner, B. M., & Dingus, T. A. (2011). Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system. *Transportation Research Part F-Traffic Psychology and Behaviour, 14*(4), 278-290. doi: 10.1016/j.trf.2011.02.001

Hellier, E., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics, 30*(2), 167-171.

Hellier, E., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35*(4), 693-706.

Ho, C., Reed, N., & Spence, C. (2007). Multisensory In-Car Warning Signals for Collision Avoidance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*(6), 1107-1114. doi: 10.1518/001872007x249965

Ho, C., Santangelo, V., & Spence, C. (2009). Multisensory warning signals: When spatial correspondence matters. *Experimental Brain Research, 195*(2), 261-272. doi: 10.1007/s00221-009-1778-5

Hollander, T. D., & Wogalter, M. S. (2000, August 2000). *Connoted hazard of voiced warning signal words: An examination of auditory components.* Paper presented at the International Ergonomics and the Human Factors and Ergonomics Society Joint Congress 2000, San Diego, CA.

Kiefer, R., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., & Shulman, M. (1999). Development and Validation of Functional Definitions and Evaluation Procedures for Collision Warning/Avoidance Systems. 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. doi: 10.1518/001872007x230271

Laughery, K. R. (2006). Safety communications: Warnings. *Applied Ergonomics, 37*(4), 467-478. doi: 10.1016/j.apergo.2006.04.020

Lee, J. D., McGehee, D. V., Brown, T. L., & Marshall, D. C. (2006). Effects of adaptive cruise control and alert modality on driver performance. *Transportation Research Record, 1980*, 49-56.

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. doi: 10.1080/00140130701318749

Lewis, B. A., & Baldwin, C. L. (2012). Equating Perceived Urgency Across Auditory, Visual, and Tactile Signals. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1307–1311.

Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for In-Vehicle Information Systems: Annoyance, Urgency, and Appropriateness. *Human Factors, 49*(1), 145-157.

Mohebbi, R., Gray, R., & Tan, H. Z. (2009). Driver Reaction Time to Tactile and Auditory Rear-End Collision Warnings While Talking on a Cell Phone. *Human Factors, 51*(1), 102-110. doi: 10.1177/0018720809333517

Neale, V. L., Perez, M. A., Lee, S. E., & Doerzaph, Z. R. (2007). Investigation of driver-infrastructure and driver-vehicle interfaces for an intersection violation warning system. *Journal of Intelligent Transportation Systems, 11*(3), 133-142. doi: 10.1080/15472450701410437

Patterson, R. D. (1982). Guidelines for Auditory Warning Systems. London: UK: Civil Aviation Authority.

Patterson, R. D. (1990). Auditory warning sounds in the work environment. In D. E. Broadbent, J. T. Reason & A. D. Baddeley (Eds.), *Human factors in hazardous situations.* (pp. 37-44): Clarendon Press/Oxford University Press.

Patterson, R., & Mayfield, T. (1990). Auditory warning sounds in the work environment [and Discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 485–492.

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. doi: 10.1518/00187200sx250674

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*(3), 153-181. doi: 10.1037/h0046162

Stevens, S. (1971). Issues in psychophysical measurement. *Psychological review*, *78*(5), 426.

Tan, A. K., & Lerner, N. D. (1995). Multiple attribute evaluation of auditory warning signals for in-vehicle crash avoidance warning systems (Report No. DOT-HS-808-535). Washington, DC.: U.S. Department of Transportation.

Teghtsoonian, M., Teghtsoonian, R., & DeCarlo, L. T. (2008). The influence of trial-to-trial recalibration on sequential effects in cross-modality matching. *Psychological Research-Psychologische Forschung, 72*(2), 115-122. doi: 10.1007/s00426-006-0103-1

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Stevens' law. *The American Journal of Psychology, 86*(1), 3-27. doi: 10.2307/1421845

Teghtsoonian, R. (2012). The Standard Model for Perceived Magnitude: A Framework for (Almost) Everything Known About It. *American Journal of Psychology, 125*(2), 165-174.

Van Erp, J. B. F., & Van Veen, H. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F-Traffic Psychology and Behaviour, 7*(4-5), 247-256. doi: 10.1016/j.trf.2004.09.003

Wiese, E. E., & Lee, J. D. (2004). Auditory alerts for in-vehicle information systems: The effects of temporal conflict and sound parameters on driver attitudes and performance. *Ergonomics, 47*(9), 965-986.

Wogalter, M. S., Conzola, V. C., & Smith-Jackson, T. L. (2002). Research-based guidelines for warning design and evaluation. *Applied Ergonomics, 33*(3), 219-230. doi: 10.1016/s0003-6870(02)00009-1

Wogalter, M. S., Kalsher, M. J., Frederick, L. J., Magurno, A. B., & Brewster, B. M. (1998). Hazard level perceptions of warning components and configurations. *International Journal of Cognitive Ergonomics, 2*(1-2), 123-143.

Wogalter, M. S., & Silver, N. C. (1990). Arousal strength of signal words. *Forensic Reports, 3*(407-420).

Wogalter, M. S., & Silver, N. C. (1995). Arousal strength of signal words: connoted strength and understandability by children, elders, and non-native English speakers. *Ergonomics, 34*, 79-89.

2.2 Validation of Essential Acoustic Parameters for Highly Urgent In-Vehicle Collision Warnings: Study 2 Overview

Study 2 was designed to validate the findings of Study 1 (and others done in the GMU Auditory Research Group Lab), in a higher fidelity context. Specifically, Study 2 assessed behavioral responses to extremely hazardous potential collision events coupled with highly urgent, time-critical warnings versus warnings missing key parameters established in prior research.

2.2.1 Validation of Essential Acoustic Parameters for Highly Urgent In-Vehicle Collision Warnings

Lewis, B. A., Eisert, J. L., & Baldwin, C. L. (2017). Validation of Essential Acoustic Parameters for Highly Urgent In-Vehicle Collision Warnings. *Human Factors*, 0018720817742114. https://doi.org/10.1177/0018720817742114

2.2.2 Abstract

Objective: Validate the importance of key acoustic criteria for use as in-vehicle forward collision warning systems.

Background: Despite recent advances in vehicle safety, automobile crashes remain one of the leading causes of death. As automation allows for more control of non-critical functions by the vehicle, the potential for disengagement and distraction from the driving task also increases. It is, therefore, as important as ever that in-vehicle safety-critical interfaces are intuitive and unambiguous, promoting effective collision avoidance responses upon first exposure even under divided attention conditions.

Methods: The current study used a driving simulator to assess the effectiveness of two warnings, one which met all essential acoustic parameters, one which only met some essential parameters, and a no warning control in the context of a lead-vehicle following task in conjunction with a cognitive distractor task and collision event.

Results: Participants receiving an FCW comprised of five essential acoustic components had improved collision avoidance responses relative a no warning condition and an FCW missing essential elements on their first exposure. Responses to a consistently good warning (GMU Prime) improved with subsequent exposures whereas continued exposure to the less optimal FCW (GMU Sub-Prime) resulted in poorer performance even relative to receiving no warning at all.

Conclusions: This study provides support for previous warning design studies and for the validity of five key acoustic parameters essential for the design of effective in-vehicle FCWs.

Application: Results from this study have implications for the design of auditory FCWs and in-vehicle display design.

Keywords: Forward collision warnings, acoustic parameters, in-vehicle warnings, auditory warnings, auditory displays, collision avoidance

Précis: The current study investigated the use of warnings adhering to inclusion of key acoustic parameters described in previous work in the context of a high-fidelity driving scenario in a motion-base simulator. Results indicate that warnings meeting five key

acoustic parameters are more effective, both during first exposure and over the course of multiple drives.

2.2.3 Introduction

Automobile crashes remain one of the leading causes of death, particularly for those under the age of 45 (Centers for Disease Control and Prevention, 2015), despite increased safety in the form of physical (e.g., seat belts, improved crashworthiness of the vehicle, etc.) and technological advances. Frontal collisions comprise a high percentage of the total crashes (Lee et. al., 2002, Scott and Gray, 2008) and therefore forward collision warning systems (FCWs) have great potential to improve safety. Simulator studies demonstrate that FCWs can decease both the severity and rate of occurrence of crashes (Baldwin, May, and Parasuraman, 2014; Brown, Lee, McGehee, 2001). Advances in sensor technologies that are capable of detecting imminent collision situations have the potential to further improve safety, but only if they can effectively communicate safety-critical information to the driver. In-vehicle safety systems and, in particular, collision warning systems, must be designed to communicate information intuitively and unambiguously in order to assist drivers with safety-critical functions and avoid increases in workload or confusion (Wiese & Lee, 2001). As automation increasingly allows more control of non-critical functions to be relinquished to the vehicle, drivers may be even more distracted compounding the need for intuitive, safety-critical interfaces.

Despite the prevalence of collisions and their impact on overall safety, collisions remain relatively rare occurrences in terms of the number of miles driven. Therefore, it is essential that FCWs be intuitively perceived as highly urgent upon first exposure. Many previous investigations of FCWs have utilized repeated collision events that may undermine the applicability of their results to naturalistic driving (Aust, Engstrom, and Vistrom, 2013).

FCWs having an acoustic component demonstrate numerous advantages over FCWs in other modalities (Dingus, Jahns, Horowitz, and Knipling, 1998; Spence and Ho, 2008). Driving is primarily a visual-manual task and therefore auditory warnings are more likely to be processed without relying on already overburdened processing resources (Wickens, 2002). Further, auditory warnings can redirect a driver's attention regardless of where the driver is looking (Baldwin, 2011). Manipulation of auditory parameters can be effectively used to impact perceived urgency across a wide range, including construction of time critical highly urgent warnings (Baldwin and Lewis, 2014). Further, simulator studies have demonstrated that collision warnings having an auditory component result in faster brake response time than visual warnings and are equally as effective as tactile warnings (Scott and Gray, 2008) even under divided attention conditions (Lewis, Penaranda, Roberts, and Baldwin, 2013). The primary goal of the current investigation was to examine the impact of key acoustic parameters on collision avoidance response the first time a driver encounters a FCW.

Previous studies have examined the impact of different acoustic parameters on an auditory signal's perceived urgency both within (Baldwin, 2011; Baldwin et al., 2012; Baldwin & Lewis, 2014) and outside the context of driving (Edworthy, 1998; Hellier & Edworthy, 1999). Increases in fundamental frequency, the presence of harmonics and rapid temporal changes increase the perceived urgency of sounds. Alarm response, as a function of differing acoustic parameters has also received considerable attention, though most frequently either outside the context of driving (Bliss, Gilson, & Deaton, 1995; Dehais et al., 2014; Edworthy & Hellier, 2006; Patterson, 1982, 1990) or after multiple exposures to repeated events (Baldwin & Lewis, 2014; Bliss & Acton, 2003; Graham, 1999; Gray, 2011; Ho & Spence, 2005). These studies have yielded important insights into the impact of different acoustic parameters on alarm response. However, the question remains: What is the impact of differing acoustic parameters of an auditory warning on collision avoidance response during first exposure?

In pursuit of this goal, research in our labs has focused on the design of sounds that are unambiguous and intuitive for a variety of driver-vehicle interface (DVI) functions. Matching the urgency of the sound to the situation it is designed to represent is one key component of effective design called Urgency Mapping (Hellier & Edworthy, 1999). Towards this end large-scale research efforts have investigated the use of psychophysical scaling to determine the inherent urgency of physical parameters of acoustic (and sometimes visual or tactile) signals (see Baldwin & Lewis, 2014; Gonzalez, Lewis, Roberts, Pratt, & Baldwin, 2012; Hellier & Edworthy, 1999; Hellier, Edworthy, & Dennis, 1993; Lewis & Baldwin, 2012; Lewis, Eisert, & Baldwin, 2014; Marshall, Lee, & Austria, 2007; Pratt et al., 2012). One particular method of psychophysical scaling is called interval scaling, and involves categorizing signals into groups or categories (Hellier, Edworthy, & Dennis, 1995), this method of psychophysical scaling has much in common with the well-explored usability practice called card sorting (see Block, Buss, Block, & Gjerde, 1981; Bonebright, Miner, Goldsmith, & Caudell, 2005; Viswanathan, Johnson, & Sudman, 1999). In a previous set of studies (Lewis, Eisert, Roberts, &

Baldwin, 2014), participants judged perceptions of urgency as directly related to acoustic characteristics for auditory DVI signals using an interval scaling paradigm displayed as a sorting task. In three separate studies, these judgments were used to predict categorization of sounds into three representative groups (Alarms, vehicle Status Notifications, and Social Notifications) based on the acoustic characteristics of the sounds in a combination of backwards and stepwise regression analyses. Five main acoustic characteristics accounted for the majority of the variance in alarm categorization: peak-to-total time ratio (or the amount of time a pulse is played at its peak intensity compared to the amount of time the pulse is played in total), inter-burst interval, the presence of harmonics, its base frequency, or lowest spectral frequency present, and pulse duration. Using the experimental data, essential acoustic criteria and "cutoffs" for acoustic parameters influencing categorization of a sound as an alarm were determined, these criteria and their cutoffs are presented in Table 6.

Table 6. Essential acoustic criteria and cutoffs determined in previous research.

Criteria	Cutoff
1. Peak to Total Time Ratio	≥ .70
2. Interburst Interval (IBI)	≤ 125 ms
3. Number of Harmonics	≥ 3
4. Base Frequency	≥ 1000 Hz
5. Pulse Duration	≥ 200 ms

These acoustic parameters were then partially validated in a subsequent experiment in which participants categorized the signals while engaged in a desktop driving simulation. Participants used vehicle controls (e.g., blinkers and brakes) to categorize. This allowed determination of both category membership and speed of categorization. Drivers responded most quickly to signals that were "unambiguous"- those which met all or none of the five primary criteria, while responses were significantly slower to ambiguous signals which met only some of the criteria.

The Current Study

The aim of current study was to validate the critical acoustic criteria when presented in potential collision situations (rather than just during categorization). Specifically, we sought to examine collision avoidance responses to a first-time exposure to signals warning drivers of a time critical hazard. A secondary aim was to examine collision avoidance response as a function of FCW acoustic parameters after exposure to a consistent (the same) or inconsistent (different) FCW while maintaining a low overall hazard event rate (Aust, et. al., 2013). Using a high fidelity driving simulator participants completed three experimental drives each of which included a difficult to avoid potential collision scenario. Though we were primarily interested in the first exposure, inclusion of two additional potential collisions allowed examination of warning consistency. Participants were divided into 5 groups with two groups encountering a consistent "optimal" versus suboptimal warning (as indicated by inclusion or exclusion of specific warning-related criteria) and two groups encountering inconsistent warnings and the fifth "control" group encountering the collision events without a warning. It was hypothesized that any warning would produce better collision avoidance responses relative to no warning (Spence and Ho, 2008), but that the warning meeting all key criteria (based on our previous research) would produce more effective collision avoidance responses relative to the warning meeting only some key criteria. Key criteria identified in our previous investigations (Lewis et al., 2014) included tempo, onset and offset ratio to total pulse time, sound frequency, the presence of multiple harmonics and, determined in subsequent research, pulse duration. It was further hypothesized that participants receiving consistent warnings would show more effective collision avoidance responses in the third collision scenario relative to participants receiving inconsistent warnings.

2.2.4 Methods

Participants

Participants were 101 undergraduate and graduate students (29 male, average age = 20.22 years) recruited through the George Mason University subject pool who volunteered in exchange for a small amount of research participation credit that could be applied to their classes. All participants self-reported normal or corrected-to-normal vision and hearing and were licensed drivers. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at George Mason University. Informed consent was obtained from each participant prior to participation in this study.

Stimuli

Stimuli consisted of two auditory warnings, both designed to have relatively high urgency and probability of being classified as a time-critical alarm, though one warning (GMU Prime) which met all five key design criteria was expected to result in somewhat better collision avoidance response than the other (GMU Sub-Prime) which met only 3 of the 5 criteria. The choice for the criteria adjusted for GMU Sub-Prime came from a previous study in our labs, investigating the effect of warnings that were each missing one of the criteria generated by Lewis et al. (2014). Lewis, Eisert, Baldwin, Singer & Lerner (2017) found that warnings that did not meet either the frequency or pulse duration criteria were less effective relative to those that met all criteria. Therefore, we chose to adjust both criteria in our GMU Sub-Prime. Both warnings were played at the same intensity level (approximating 10 dB above ambient background engine noise) while the simulator was running. An additional 29 other sounds, including alarms, status notification sounds and social notification sounds previously examined by our labs, were also presented. Specific parameters that were different for the two warnings were pulse duration (GMU Prime: 200 ms or GMU Sub-Prime:400 ms), base frequency (GMU Prime: 1576 Hz or GMU Sub-Prime: 3000 Hz) and peak to total time ratio (GMU Prime: .95 or GMU Sub-Prime: .9, both of which are within the appropriate criteria, but differed slightly as an effect of the difference in pulse duration while holding onset and offset constant). Warnings played for a duration of 1600 to 2200 ms, had onset and offset times of 10 ms, and multiple harmonics. Due to the presence of 10 ms onset and offset times the perceived interpulse interval (IPI) was approximately 18 ms. The perceived IPI is based on standards established for medical alerts that define downtime or perceived alert off time to consist of any part of the sound below 90% intensity (International Electrotechnical Commission, 2006). Figure 9 shows the perceived IPI of 18 ms, created by the 10 ms onset and offset times.

Figure 9. Screenshot of two GMU Prime pulse waveforms and the onset and offsets that created the perceived interpulse interval time.

Participants were randomly assigned to one of five groups. Two groups received consistent warnings, either all GMU Prime or all GMU Sub-Prime. Two groups included “switch” conditions, where the warning they received was inconsistent throughout the course of three experimental drive events (eg, GMU Prime, then GMU Sub-Prime, then GMU Prime), and one group, the control group, received no warnings during any of the experimental drive events. Groups and warnings are elaborated in Table 7.

Table 7. Group warning characteristics by drive

Group	Drive 1 Event	Drive 2 Event	Drive 3 Event
GMU Prime	GMU Prime	GMU Prime	GMU Prime
GMU Sub-Prime	GMU Sub-Prime	GMU Sub-Prime	GMU Sub-Prime
GMU Prime Switch	GMU Prime	GMU Sub-Prime	GMU Prime
GMU Sub-Prime Switch	GMU Sub-Prime	GMU Prime	GMU Sub-Prime
No Warning	No Warning	No Warning	No Warning

Further, following first exposure to the collision event, the second and third drives exposed participants to a task where they were asked to categorize twenty-nine additional sounds not included for use as warnings in the main study. These sounds varied on all parameters and were implemented in part to engage participants in an additional task in an effort to minimize expectancy of an additional collision event. The 29 additional sounds included currently in-use vehicle sounds (and some sounds designed by our labs) that were intended to represent less urgent notifications, such as lane deviation warnings, curve speed warnings, fatigue alerts, backup and park assist sounds, seatbelt reminders, door open reminders, and various types of infotainment and social notifications.

Apparatus and Procedure

The experiment was run in a Realtime Technologies, Inc. (RTI), open-cab driving simulator on a motion-base. The motion-base allowed for 180 degrees of yaw motion to simulate turns, and one degree of pitch motion to simulate acceleration and braking. The visual component of the simulator included three 42-inch plasma displays, allowing for a 180-degree field of view (Figure 10). An RTI program called SimVista was used to create two simulated driving worlds and all scenarios. Data were collected at 30 Hz. Prior to the experiment participants gave written informed consent and verbally completed a motion sickness history screener to assess susceptibility to simulator adaptation syndrome (SAS: see (Mollenhauer, 2004 for a review). Participants scoring over a 7 on the questionnaire were given the option to opt out of the experiment. Only two participants (both female) were unable to participate due to susceptibility.

Figure 10. Motion-base advanced driving simulator used to run driving scenarios.

After completion of the motion sickness history screener and informed consent, participants were introduced to the simulator. All participants were given basic safety instructions and were required to buckle their seatbelt in order to complete all drives. Participants completed two practice drives prior to the first experimental drive. First, participants practiced driving alone, with no secondary tasks. Participants were instructed to drive, following a lead car in front of them at a speed of 65 mph (though, due to the tightness of curves in the first driving world, participants were instructed to slow down when turning). They were instructed to, remain in the right-hand lane at all times. Participants drove until both they and the experimenter felt comfortable with their driving performance (i.e., no skidding around turns, proper lane and speed maintenance). After the first practice drive, participants were introduced to the subsidiary task, a visual-manual 1-back task. The task required participants to monitor a small touchscreen to the right of the steering wheel which constantly presented numbers from 0-9 along with the words “YES” and “NO”. Participants were required to respond by pressing the corresponding affirmative or negative button based on whether the number presented matched or did not match the number presented directly preceding the currently presented number (see Figure 11). For Drive 1, numbers would appear for 2 seconds during which time participants could respond. Responses, whether correct or incorrect were immediately followed by another stimulus.

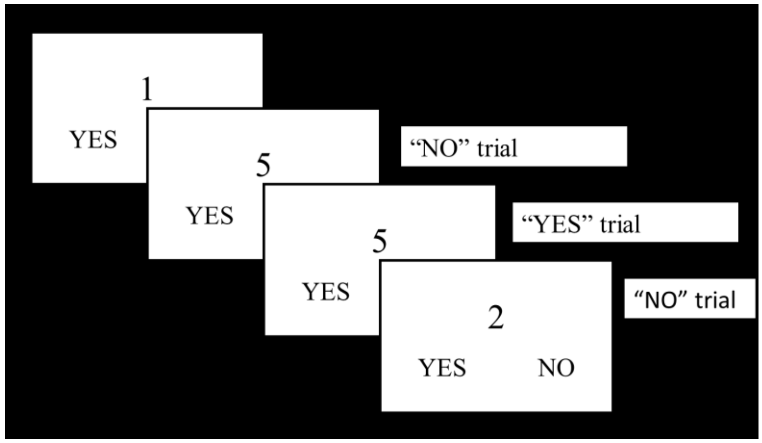

Figure 11. 1-back design and correct responses.

Once participants were comfortable completing the subsidiary task by itself, they repeated the first practice drive scenario, this time while completing the subsidiary task. All participants were instructed to prioritize driving safely, meaning that if they were uncomfortable with the subsidiary task during complex maneuvers (like during turns) they should stop doing the task, and return to it when they felt in control of the vehicle. Participants again completed the dual-task practice until the experimenter felt comfortable with their control of the vehicle during dual-task phases. Each practice took a varying amount of time depending on how long it took for the participant to reach satisfactory performance, which consisted at a minimum of drivers being able to maintain the position of the simulated vehicle within the intended lane and maintain a speed within a 5-mile range for a period of at least 2 minutes.

After completion of both practice drives, participants completed the first experimental drive. This drive was the same as the drive used in the two previous practices, seemingly the exact same as the preceding dual-task practice drive, however, about 3 minutes into the drive (after the first two turns) while driving on a straight section of roadway the lead car changed lanes suddenly, revealing an almost stopped car in front of it. At this point, participants either received one of the two warnings or received no warning based on their group. Participants could either employ a hard brake or a swerve to successfully avoid collision. However, this event was designed to be extremely difficult to avoid. The experimenter ended the after the participant either crashed or about 30 seconds after the successful avoidance of the scenario.

After the first experimental drive (each practice and drive was ended with the closing of the simulation, loading of a new scenario, and the initiation of a new simulated drive), participants were told that their vehicle was now a "connected vehicle" and given instructions for responses to sounds (from the 29 extra sounds) that they would hear. Participants were told that sounds could fall into one of three categories: alarms, status notifications and social notifications. These were defined for participants and were matched to appropriate responses where alarms should be responded to with a brake press, status notifications should be responded to by pressing a triangle indicator button that would appear in place of the secondary task and social notifications should be responded to by pressing a telephone button that would appear along with the triangle indicator on the touchscreen (Figure 12).

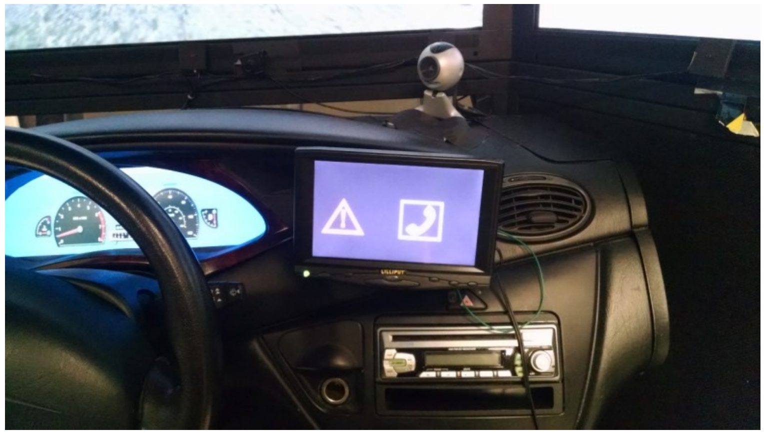

Figure 12. Touchscreen images for "Status" and "Social" categorization.

Participants were instructed that there were no right or wrong answers and that the purpose of this exercise was to collect their interpretations of the intended sound categories. After practicing these responses participants completed two more experimental drives lasting 15 and 20 minutes (respectively) where they responded to alarm, status and social notifications followed by an event. In the second experimental drive, the event was a lead vehicle braking event and in the third experimental drive the event included a reveal event, identical to the event in the first experimental drive. For Drive 2, an interstimulus interval (ISI) of between 4 and 7 seconds was added between trials on the n-back task and there were between 9 and 13 trials between sounds. For Drive 3 the ISI was adjusted to between 3 and 7 seconds and there were only between 4 and 7 trials between sounds to decrease the amount of total time needed for the drive. The entire experiment took a little under 2 hours.

Due to different possible responses and appropriate actions for the second drive (which included a different hazard event than the first and third events, which did not differ), analysis of the second drive event has been excluded from this report. The main comparison of interest is response to hazards in Drive 1 and 3, as these were the same lead-car reveal event.

Design and Data Analysis

Independent variables included warning played (GMU Prime, GMU Sub-Prime and no warning) and Group (1-5) with main dependent variables of interest being collisions, evasive maneuver response time (EMRT) and speed at collision for those participants who did collide as an index of collision severity.

Table 8. Simulator metrics and their descriptions and units

Metric	Description
Accelerator Release Time	The time from the onset of the warning to the time that the participant released the accelerator (ms)
Outcome	Whether the participant collided or avoided the collision (0 or 1)
Braking Response Time (BRT)	The time from the onset of the warning to the time that the participant touched the brakes (ms)
Distance at Warning	The distance from the participant's vehicle to the revealed car at the time of the warning (meters)
Speed at Collision	The speed at which the participant collided, if the participant collided (mph)
Headway Time at Warning	The participant's headway time to the revealed car at the time of the warning (ms)
Max Brake Force	The max force that the participant applied to the brakes (N)
Minimum Distance	The minimum distance that the participant reached in relation to the reveal car (meters)

Minimum Headway Time	The minimum headway time that the participant reached in relation to the revealed car (ms)
Evasive Maneuver Response Time (EMRT)	The time from the onset of the warning until any evasive maneuver by the participant including braking or a swerve response (ms). Swerve response was determined by taking the derivative of their steering input and any value greater than 1.5 was considered a swerve response.
Speed at Warning	The speed at which the participant was travelling at the time of the warning (mph)
Speed Reduction	The amount by which the participant reduced their speed in total (mph)
Time from Initial Response to Max Brake Force	The time from the first brake response by the participant to the time that they applied their maximum brake force (ms)
Time to Brake from Accelerator Release	The time from the release of the accelerator by the participant to the time that they engaged the brakes (ms)
Time to Max Brake Force	The time from the onset of the warning to the time that the participant reached their maximum brake force (ms)

The full list of metrics and their descriptions taken from the simulator are listed in Table 8. Figures were created in Microsoft Excel and all error bars indicate +/- standard error values.

2.2.5 Results

Table 9 gives a breakdown of participant demographics by group. Effort was made to ensure equal gender and age distributions across groups.

Table 9. Breakdown of participant demographics by group

Group	Male	Female	Age	*SD* (age)	Total
GMU Prime	5	15	19.55	1.67	20
GMU Sub-Prime	6	14	22.30	6.30	20
GMU Prime Switch	7	12	19.47	1.93	19
GMU Sub-Prime Switch	5	15	19.45	2.54	20
No Warning	6	16	20.32	2.64	22
Total	29	72	20.23	3.55	101

Analysis of demographic data indicates that there were no significant differences in gender distribution by group: $F(4,96) = .197$, $p = .939$, and only a marginal difference between age between groups, $F(4,96) = 2480$, $p =.05$, though age was not a significant predictor of collisions or speed at time of collision, $F(12,88) = .680$, $p = .767$ and $F(8,47) = .569$, $p = .798$, respectively.

Drive 1

The primary aim of this paper was to examine collision avoidance response upon first exposure to a collision event. Results for Drive 1 were analyzed in terms of warning played rather than by group, as Groups 1 and 3 and Groups 2 and 4 received identical alerts up until Drive 2. Analysis of collisions by groups indicate that participants who received a warning collided somewhat (though not significantly) less often than did participants who received no warning (collapsing warning groups versus no warning, Table 10).

Table 10. Collisions and avoidances by warning played

Warning Played	Avoided	Collided	Total
GMU Prime	19 (47.5%)	21 (52.5%)	40
GMU Sub-Prime	19 (47.5%)	21 (52.5%)	40
No Warning	7 (33.3%)	14 (66.6%)	21

Further analysis of data for participants who collided indicated that, although there were not statistically significant differences in number of collisions between groups, there were significant differences in speed at time of collision, $F(2,53) = 4.01$, $p = .024$. This metric represents the speed of the participant's vehicle at the time that it collided with the stopped reveal car and can be considered as a metric of collision severity. Figure 1313 shows that those who received a GMU Prime were traveling at significantly reduced speed upon impact relative to those who did not receive a warning. Post hoc comparisons using the Tukey HSD test indicate that only the difference between GMU Prime and no warning was significant ($p = .019$), where Warning 2 did not vary significantly from either group.

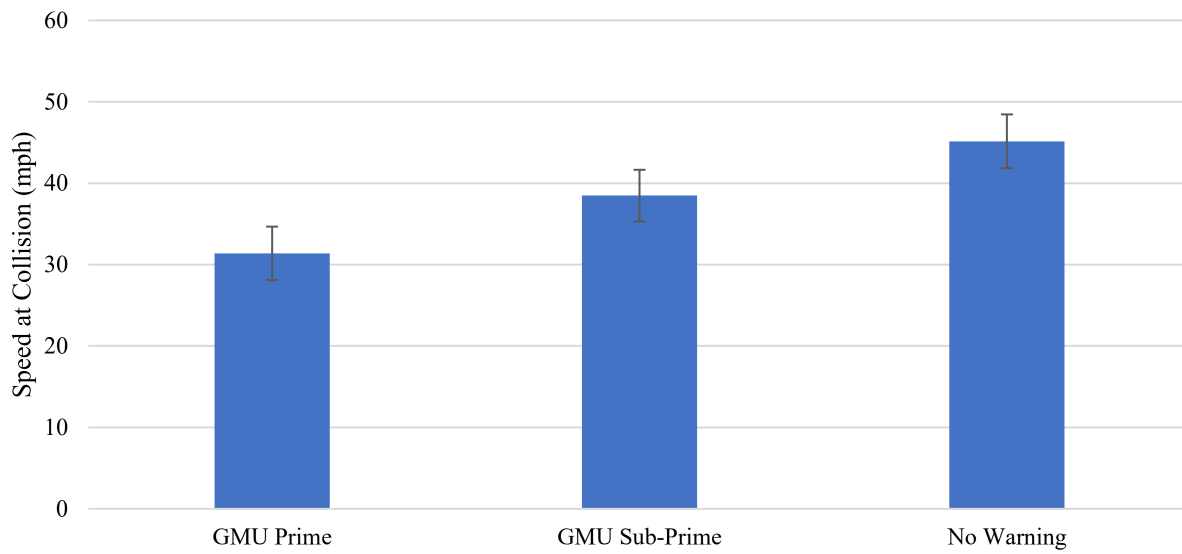

Figure 13. Drive 1 speed at time of collisions (if participant collided) by warning.

Further, results indicate that brake response time (BRT), though not emergency maneuver response time, varied significantly by group, $F(2,66) = 3.28$, $p = .044$. We chose to assess both BRT and EMRT because a participant's first response could have been either to brake or steer. Both responses are derived from the simulator output: BRT is defined as the time from the event onset to the first detectable pressure on the brake pedal, whereas EMRT is the time from the event onset until either detectable brake pressure or a detectable wheel response (commonly referred to as "swerving"), allowing us to account for participants who did not brake to the event but merely steered around the stopped vehicle. Figure 14 shows differences in BRT by warning. Tukey HSD post hoc comparisons indicate that those receiving GMU Prime had significantly faster BRT relative to the no warning group, (p=.050), a difference of about 350 ms, and that the GMU Sub-Prime group did not differ from either GMU Prime or the No Warning control group.

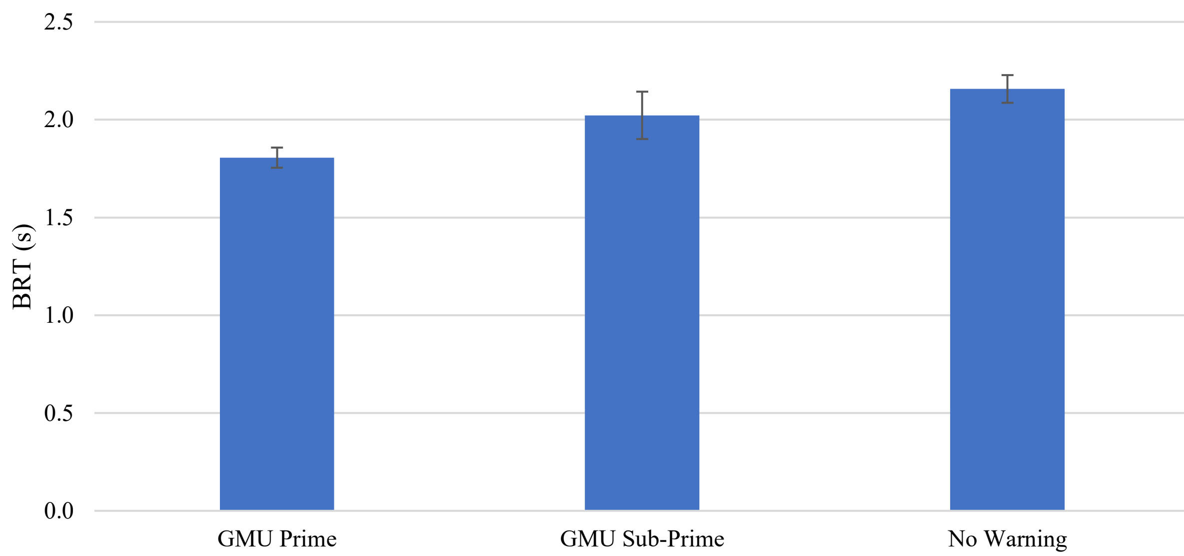

Figure 14. Brake response time (BRT) by warning.

It is important to point out that this event happened very early into the experiment, after adequate but short practice with the simulator. All participants included in this experiment indicated that they had no previous experience using a motion based driving simulator. Additionally, the event itself was designed to be representative of extremely difficult collision events and therefore the very high collision rate overall is not unexpected. Importantly, despite the high collision rate for all groups, results indicate that participants receiving a warning that met all 5 design criteria (GMU Prime) collided around 15 mph slower than those who received no warning.

Drive 3

Analysis of drive 3 collision data indicates no significant differences in speed at collision for participants who collided, $F(4,25) = 1.55$, $p = .217$. However, it was found that there was a homogeneity of variance (as assessed by Levene's Equality of Variances Test) therefore we additionally conducted an independent samples t-test was conducted. Results indicate that there were significant differences between the GMU Prime group (M = 14.7 mph, SD = 8.3 mph) and the GMU Sub-Prime group (M = 42.3 mph, SD = 18.5 mph); $t(11) = -2.46$, $p = .032$. The few participants in the consistent GMU Prime condition who did collide did so at a lower speed than did participants in all other conditions (Figure 15).

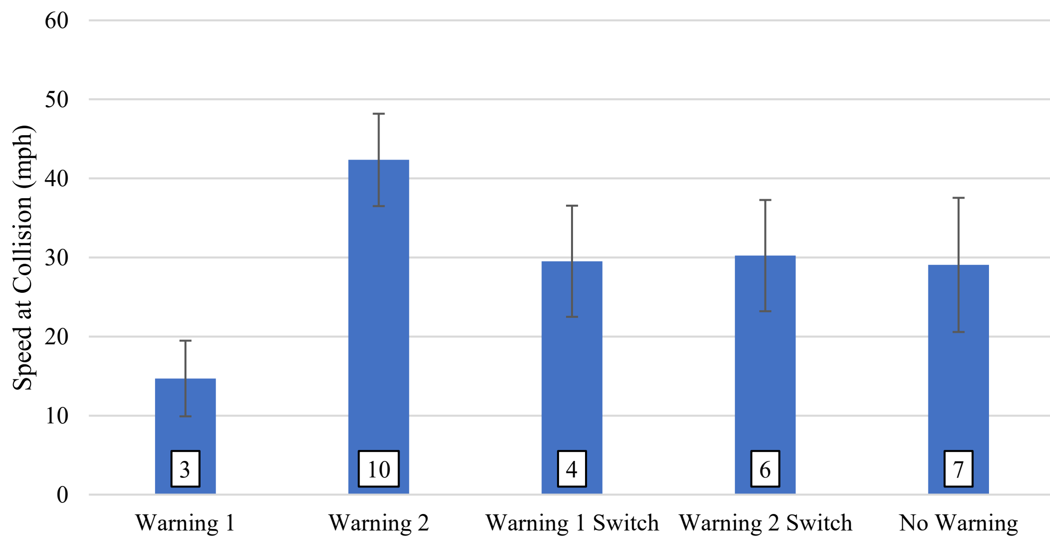

Figure 15. Speed at time of collision for participants who collided by group, values inside base end represent the number of collisions, error bars represent standard error

No other dependent variables revealed significant differences between groups.

2.2.6 Discussion

Results from the current study provide validation for the importance of the five key acoustic criteria described here when designing time critical FCWs. Specifically, current results indicate that ensuring that an auditory warning meets the five essential acoustic parameters will increase the likelihood that drivers engage in effective collision avoidance responses when presented with a potentially imminent forward collision event, even when that event is unexpectedly encountered for the first time. The absence of two or more of these key parameters results in ambiguity that decreases effective collision avoidance response. Further the present results support findings from previous studies in our lab that these five key parameters result in more effective collision avoidance performance relative to warnings meeting only some criteria upon subsequent exposures (Lewis, Eisert, Roberts & Baldwin, 2014). For the participant's first exposure, we found evidence similar to previous investigations that any warning is better than no warning, in terms of decreased response times, decreased collision rates and trends toward decreased speed at the time of collision if a collision occurs, compared to participants receiving no warning. The significant effect of speed at time of collision indicated that only GMU Prime resulted in significantly reduced speed (a proxy for crash severity) relative to the no warning condition on first exposure. GMU Sub-Prime, a clearly audible auditory warning, was not able to achieve this safety benefit. Further, the current results extend this by showing that the most effective warning across multiple exposures - including the first - was GMU Prime, the one constructed using all five key acoustic parameters. Less effective warnings or inconsistent warnings do not appear to have an aiding effect on collision avoidance response after repeat exposures, though results are not statistically significant. Participants who consistently received only GMU Prime had the most effective collision avoidance responses compared to other groups in terms of both reduced collisions and reduced speed at time of collision. Participants consistently receiving GMU Sub-Prime, the warning missing key components, performed the worst of all groups. This includes the apparent fact that participants consistently receiving a poor warning or inconsistent warnings did not seem to show the same improvement, as reflected by no changes in speed reduction at the time of collision for participants who collided, though findings were again inconsistent. Conversely, participants in the GMU Prime and no warning condition did show decreases in collision severity by Drive 3. We predicted more effective collision avoidance response for consistent presentation of GMU Prime, relative to GMU Sub-Prime. However, we did not expect consistent presentation of GMU Sub-Prime (a less optimal warning) to negatively impacted performance. We do not have a conclusive explanation for this result. It is possible that the difficulty of the second collision event affected trust in the forward collision warning, specifically for the group consistently receiving GMU Sub-Prime. The hazard event in Drive 2 was nearly impossible to avoid. This may have caused participants to mistrust the warning they received in that drive. The GMU Prime Group might have mistrusted GMU Prime, but it was a good enough warning that they still responded appropriately. The switch groups received different warnings for Drive 2 and then the same warning that helped them in Drive 1 was present again in Drive 3 causing them to react similarly. for the GMU Sub-Prime Group, their warning may have been perceived as less urgent (as indicated by Drive 1 data), and in Drive 2 it may have been perceived as a malfunction as it was unable to improve performance. Therefore, in the final drive participants in the GMU Sub-Prime Group may have been less likely to trust the warning system. It should also be noted that we are unable to quantify the effect of the presence of additional sounds to which participants were required to respond in Drives 2 and 3 as we did not have a group that had no additional alerts. It is possible that the presence of many sounds to which a response was required but that had no obvious cause (these alerts were not linked to any event) changed the way that participants responded to the collision scenarios and warnings.

Findings from this study support the use of criteria defined in (Lewis, Eisert, Roberts, & Baldwin, 2014), for DVI design with the caveat that safety-critical signals should meet all five key acoustic parameters in order to ensure that information is unambiguous and intuitive. It must also be noted that this study has limitations. First, this study used a motion-base driving simulator rather than more naturalistic driving conditions for obvious safety reasons. Therefore, the generalizability to more realistic driving situations can be questioned. However, in addition to safety, the use of simulation allowed for the potential collision events in Drive 1 and Drive 3 to be identical in timing, traffic present and abilities of the simulated vehicle. Generalizability for this study may also be affected by the relative youth of the sample population, both in terms of their response times (as increased age typically corresponds to a decrease in response times) and their experience driving (this sample reported an average of 3.5 years of driving). An additional limitation is that in the current investigation the warnings were presented alone, as opposed to over the backdrop of music or conversation. Although research has attempted to identify whether there are likely to be unintended or unforeseen interactions between the recommended signals and ambient noise (see Lerner, Singer, Kellman & Traube, 2015), and warnings meeting the five key criteria recommended here appear to be both detectable and recognizable in naturalistic driving conditions with more realistic ambient noise (e.g., while listening to music or driving with the windows rolled down), more research in this area is needed. It is likely that participants also behaved differently in the simulator than they would in normal traffic. In particular, participants were not allowed to have cell phones in the cab of the simulator and they were aware that they were being observed at all times. This may have contributed to overly safe driving, as well as given advance indication that a collision or surprise event might occur. It must also be pointed out that the current investigation did not manipulate signal intensity. For the current purposes, which focused on recognition rather than detection, this important acoustic characteristic was held constant across two warnings. Fluctuating ambient background noise levels will impact alert detection and it will be essential that FCWs are played at appropriately high levels, regardless of their other characteristics. Finally, it was only possible to evaluate two types of collision warnings due to the nature of the simulated experiment and the desire to maintain a low collision rate (Aust, et. al., 2013). It is recommended that future research explore variations in non-essential acoustic parameters while adhering to set criteria for essential parameters as well as investigate possible interactions with annoyance or familiarity, varying ambient background noise, as well as false alarms over time.

In summary, auditory FCWs can be designed to be intuitive and effective upon first exposure to an unexpected, highly critical, potential collision event and to remain effective upon subsequent repeated exposures. However, it is essential that the auditory FCW be comprised of the following five key elements. Specifically, the auditory FCW should have a peak-to-total time ratio of greater than .7, and inter-burst interval of less than or equal to 125 ms, at least three harmonic components, a base frequency of at least 1000 Hz, but no more than 2500 Hz, and a pulse duration of 200 ms or less. Further research to validate these essential acoustic parameters in more naturalistic driving conditions is clearly warranted. It will be important for future research to investigate the effects of using these types of highly urgent alarms in automated or autonomous systems. Specifically, how acceptable alerts with these parameters be in unreliable systems where there are many false alarms? Would these alerts be appropriate in vehicles with advanced or automatic braking systems, where the alert is intended to warn the operator that a hard brake is going to occur as opposed to warning the driver that they must initiate a hard brake? However, the converging evidence for the effectiveness of these parameters across multiple methodologies in our previous studies and their beneficial impact on improving collision avoidance response even under divided attention conditions and unexpected first exposure demonstrate strong support their importance as essential acoustic parameters for constructing effective high urgency auditory FCWs.

2.2.7 Key Points

- This study replicated and validated the findings from work done previously investigating appropriate acoustic parameters and criteria for warning design.
- Warnings which adhere to the criteria, set forth in this study, for good warnings help drivers to respond more appropriately in high-urgency scenarios.
- Drivers receiving inconsistent warning showed smaller improvements over time and sometimes decrements in performance relative to those receiving consistent good warnings or even no warning.

2.2.8 References

Aust, M. L., Engström, J., & Viström, M. (2013). Effects of forward collision warning and repeated event exposure on emergency braking. *Transportation Research Part F: Traffic Psychology and Behaviour, 18*, 34-46. doi: 10.1016/j.trf.2012.12.010

Baldwin, C. L. (2011). Verbal collision avoidance messages during simulated driving: Perceived urgency, alerting effectiveness and annoyance. *Ergonomics, 54*(4), 328-337. doi: 10.1080/00140139.2011.558634

Baldwin, C. L., Eisert, J. L., Garcia, A., Lewis, B., Pratt, S. M., & Gonzalez, C. (2012). Multimodal urgency coding: Auditory, visual, and tactile parmeters and their impact on perceived urgency. *Work, 41*, 3586-3591. doi: 10.3233/WOR-2012-0669-3586

Baldwin, C. L., & Lewis, B. A. (2014). Perceived urgency mapping across modalities within a driving context. *Applied Ergonomics*, *45*(5), 1270–1277.

Baldwin, C. L., May, J. F., & Parasuraman, R. (2014). Motor Vehicle Crashes Due to Task-Induced Fatigue in Young and Older Drivers: Mitigation with Auditory Forward Collision Warnings. *International Journal of Human Factors and Ergonomics, 135*(4), 2280-2286.

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

Block, J., Buss, D. M., Block, J. H., & Gjerde, P. F. (1981). The cognitive style of breadth of categorization: Longitudinal consistency of personality correlates. *Journal of Personality and Social Psychology*, *40*(4), 770–779.

Bonebright, T. L., Miner, N. E., Goldsmith, T. E., & Caudell, T. P. (2005). Data collection and analysis techniques for evaluating the perceptual qualities of auditory stimuli. *ACM Trans. Appl. Percept.*, *2*(4), 505–516. https://doi.org/10.1145/1101530.1101550

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.

Centers for Disease Control and Prevention. (2015). *Leading causes of death reports. National and regional 1999—2014*.

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 56*(4), 631-644. doi: 10.1177/0018720813510735

Dingus, T. A., Jahns, S. K., Horowitz, A. D., & Knipling, R. (1998). Human factors design issues for crash avoidance systems. *Barfield, Woodrow*.

Edworthy, J. (1998). Warnings and hazards: An integrative approach to warnings research. *International Journal of Cognitive Ergonomics, 2*(1-2), 3-18.

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: implications for medical alarms. *BJA: British Journal of Anaesthesia, 97*(1), 12-17. doi: 10.1093/bja/ael114

Gonzalez, C., Lewis, B. A., Roberts, D. M., Pratt, S. M., & Baldwin, C. L. (2012). Perceived Urgency and Annoyance of Auditory Alerts in a Driving Context. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1684–1687.

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

Gray, R. (2011). Looming Auditory Collision Warnings for Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(1), 63-74. doi: 10.1177/0018720810397833

Hellier, E., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics*, *30*(2), 167–171.

Hellier, E., Edworthy, J., & Dennis, I. (1995). A comparison of different techniques for scaling perceived urgency. *Ergonomics*, *38*(4), 659–670.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors*, *35*(4), 693–706.

Ho, C., & Spence, C. (2005). Assessing the Effectiveness of Various Auditory Cues in Capturing a Driver's Visual Attention. *Journal of Experimental Psychology: Applied, 11*(3), 157-174.

International Electrotechnical Commission. (2006). IEC 60601-1- 8: Medical electrical equipment—General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems. Geneva: Author.

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.

Lerner, N., Singer, J., Kellman, D., & Traube, E. (2015). In-Vehicle Noise Alters the Perceived Meaning of Auditory Signals. Presented at the 8th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.

Lewis, B. A., & Baldwin, C. L. (2012). Equating Perceived Urgency Across Auditory, Visual, and Tactile Signals. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1307–1311.

Lewis, B. A., Eisert, J. L., Roberts, D. M., & Baldwin, C. L. (2014). Designing Unambiguous Auditory Crash Warning Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 2078–2082. https://doi.org/10.1177/1541931214581437

Lewis, B., Eisert, J., & Baldwin, C. (2014). Effect of Tactile Location, Pulse Duration, and Interpulse Interval on Perceived Urgency. *Transportation Research Record: Journal of the Transportation Research Board*, (2423), 10–14.

Lewis, B. A., Eisert, J. L., Baldwin, C. L., Singer, J., & Lerner, N. (2017, March). Urgency coding validations (Final report. Report No. DOT HS 812 385). Washington, DC: National Highway Traffic Safety Administration.

Lewis, B. A., Penaranda, B. N., Roberts, D. M., & Baldwin, C. L. (2013). Effectiveness of bimodal versus unimodal alerts for distracted drivers.*Human Factors in Driver Assessment, Training and Vehicle Design*, 376-382.

Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for In-Vehicle Information Systems: Annoyance, Urgency, and Appropriateness. *Human Factors*, *49*(1), 145–157.

Mollenhauer, M. A. (2004). *Simulator adaptation syndrome literature review*. DTIC Document.

Patterson, R. D. (1982). Guidelines for Auditory Warning Systems. London: UK: Civil Aviation Authority.

Patterson, R. D. (1990). Auditory warning sounds in the work environment. In D. E. Broadbent, J. T. Reason & A. D. Baddeley (Eds.), *Human factors in hazardous situations.* (pp. 37-44): Clarendon Press/Oxford University Press.

Pratt, S. M., Lewis, B. A., Peñaranda, B. N., Roberts, D. M., Gonzalez, C., & Baldwin, C. L. (2012). Perceived Urgency Scaling in Tactile Alerts. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1303–1306.

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. doi: 10.1518/00187200sx250674

Spence, C., & Ho, C. (2008). Multisensory warning signals for event perception and safe driving. *Theoretical Issues in Ergonomics Science, 9*(6), 523-554. doi: 10.1080/14639220701816765

Viswanathan, M., Johnson, M. D., & Sudman, S. (1999). Understanding consumer usage of product magnitudes through sorting tasks. *Psychology & Marketing*, *16*(8), 643–657.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wiese, E., & Lee, J. D. (2001). Effects of Multiple Auditory Alerts for in-Vehicle Information Systems on Driver Attitudes and Performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(23), 1632–1636.

2.3 Individual Differences in Intramodal and Crossmodal Inattentional Insensitivity: Study 3 Overview

The third and final study in this series was designed to improve our understanding of why some individuals may still miss highly urgent and time-critical signals despite the theoretical effectiveness of the warning. In Study 3, we assessed the effect of working memory capacity on sensitivity to intramodal or crossmodal signals in perceptually demanding auditory and visual tasks.

2.3.1 Individual Differences in Intramodal and Crossmodal Inattentional Insensitivity

2.3.2 The Current Study

It was the goal of the current study to investigate the relationship between individual differences in working memory capacity (WMC) and inattentional insensitivity to salient signals. It has been theorized that individual differences may explain some of the variability in inattentional insensitivity that have been found throughout intramodal and crossmodal, laboratory and real-world tasks (see: Beanland & Pammer, 2010) ; Macdonald & Lavie, 2011; Dalton & Fraenkel, 2012; Dattel et al., 2013; Chabris, Weinberger, Fontaine, & Simons, 2011, etc)

Previous work has been inconsistent in both methodology and findings, meaning that comparison between studies is challenging to say the least. There is evidence (often conflicting) that inattentional insensitivity is directly related to critical stimulus salience, primary task difficulty, expertise on primary task, age (broadly), physical processing abilities, WMC and cognitive control. These can, importantly, be broken up into task factors (either primary or secondary) and person factors (individual differences).

The current study investigated individual differences in inattentional insensitivity as a function of WMC. Specifically, it assessed whether inattentional insensitivity can be explained by a single-route WMC model or by a dual-route WMC model. A single-route WMC model would predict that individuals with higher WMC would show lower levels of inattentional insensitivity in intramodal inattentional insensitivity tasks (IIS) relative to individuals with low WMC. However, this theory would predict that in crossmodal inattentional insensitivity (CIS) tasks, both high and low WMC individuals would have similar insensitivity levels. Conversely, a dual route model would predict that individuals with either high levels of WMC or low levels of WMC would show higher levels of inattentional insensitivity as compared to those with middling levels of WMC in IIS tasks. The dual route model suggests that high WMC individuals may overactively inhibit attention to irrelevant stimuli and low WMC individuals may not have resources available to attend to secondary stimuli. In this case, individuals with middling levels of

WMC should show the lowest inattentional insensitivity. Additionally, according to the dual route model all individuals would be expected to show similar levels of inattentional insensitivity in CIS tasks.

The objective of the current study was to systematically investigate the dual- vs. single- route hypothesis. In order to determine which model had the most accurate fit, individuals were assessed on WMC (via the OSpan task) and inattentional insensitivity in intramodal and crossmodal paradigms. Specifically, primary tasks were perceptually demanding visual or auditory tasks and critical stimuli were presented in either the same or the opposite modality and consisted of task-irrelevant signals. Listed below are the a priori hypotheses based on the literature. Each set of hypotheses is followed by a graphic that shows the theoretic outcome graph that would be associated with support for those hypotheses.

Dual Route Model Hypotheses:

H_1: Individuals with very high WMC will show high levels of inattentional insensitivity in both crossmodal and intramodal tasks as they will actively inhibit task-irrelevant processing.

H_2: Individuals with very low WMC will show high levels of inattentional insensitivity in crossmodal and intramodal tasks as they will have limited spare resources with which to processes task-irrelevant stimuli.

H_3: Individuals with middle WMC will show the lowest levels of inattentional insensitivity across all tasks.

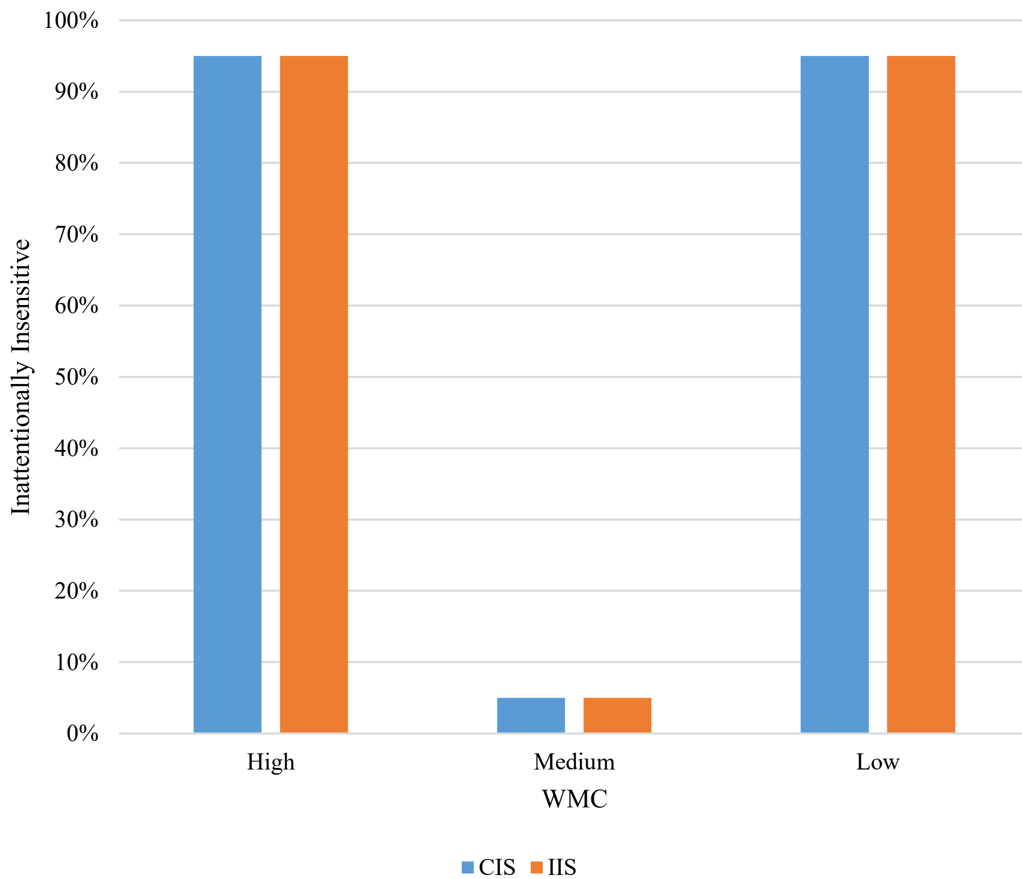

Figure 16. Theoretical graph for the hypotheses

Alternate Hypotheses

H_{1-0}: Individuals with very high WMC will have the lowest levels of inattentional insensitivity as they will have spare resources to process additional information.

$H_{2\text{-}0}$: Individuals with very low WMC will have low levels of inattentional insensitivity because they will be unable to actively inhibit task-irrelevant information.

$H_{3\text{-}0}$: Individuals in with all levels of WMC will show lower levels of inattentional insensitivity in crossmodal than in intramodal trials.

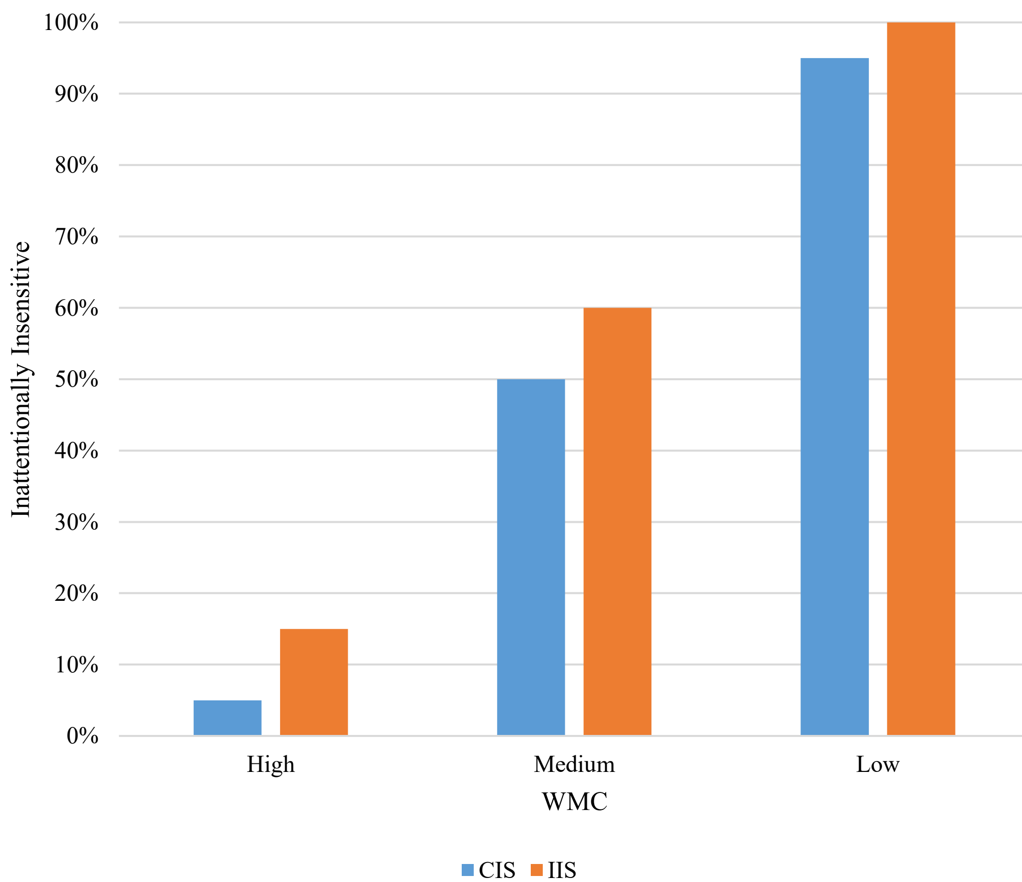

Figure 17. Theoretical graph for the alternate hypotheses

Null Hypothesis

H_0: Inattentional insensitivity will be present for a proportion of participants in all groups regardless of WMC or modality.

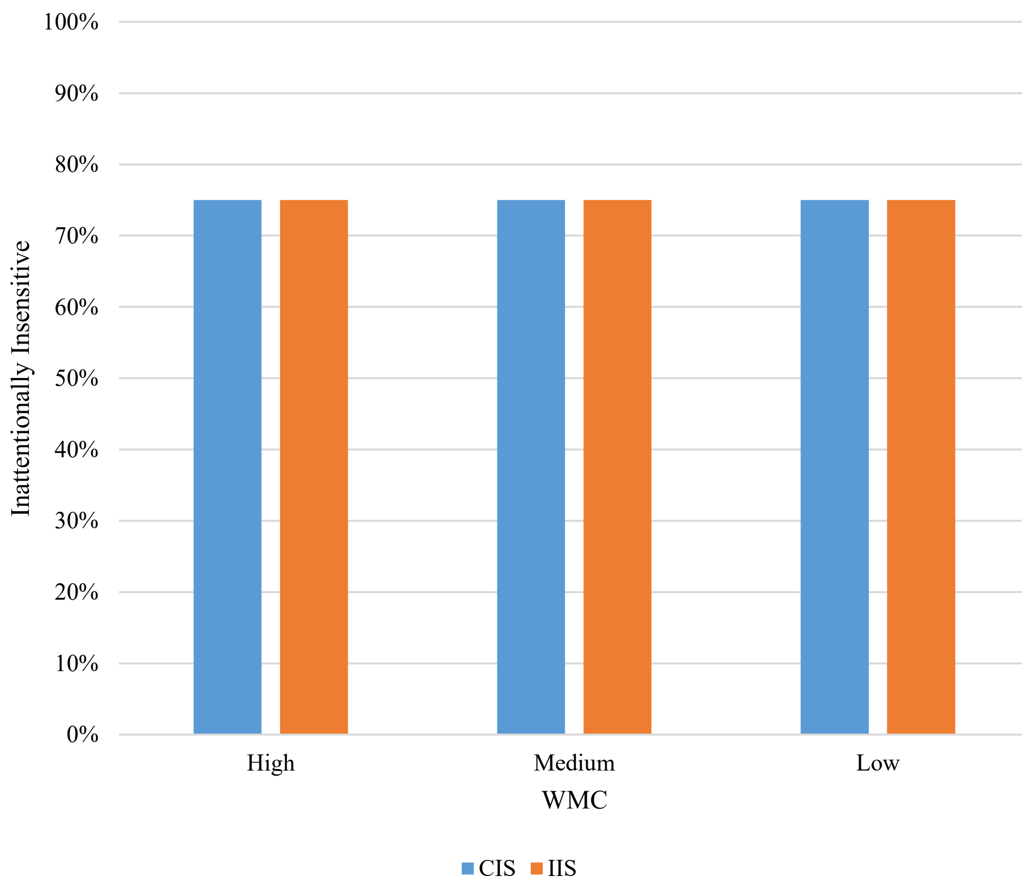

Figure 18. Theoretical outcome graph for the null hypothesis

2.3.3 Methods

Participants

Data were collected from 157 participants in total. Fourteen participants did not provide age and/or gender information. Of the remaining 143 participants, 40 were male and 103 were female, with an average age of 20.10 years (and standard deviation of 2.98 years). Participants were recruited from the George Mason University psychology research participation pool or participated on a volunteer basis.

Stimuli and Apparatus

Intramodal Inattentional Insensitivity Tasks

Visual task-visual critical stimulus

The visual experimental task was modeled after the task used by Mack & Rock (1998) and Macdonald & Lavie (2011). Subjects were shown a black fixation circle in the middle of a gray (RGB values: 204R 204G 204B) screen for 900 ms, followed by a cross display for 200 ms and then received a response window consisting of a blank, gray screen for up to 2500 ms. If a participant made a response, the task moved on immediately. The cross consisted of one blue (RGB values: 0R 183G 255B) and one green (RGB values: 0R 204G 0B) arm and one long and one short arm. Each combination was equally likely but one arm was always blue and the other green and one was always long and the other short. The cross was displayed on a gray background. The visual critical signal (CS) was presented in the first block on trial 7, and then was presented randomly in 6 of the next 8 blocks, and again for the last block (Figure 19). The visual CS was a small shape (circle, square, diamond, or triangle) presented in white with blurred edges in one of the quadrants of the cross. Shapes were equated for perceived size using a short psychophysical matching pilot test prior to data collection.

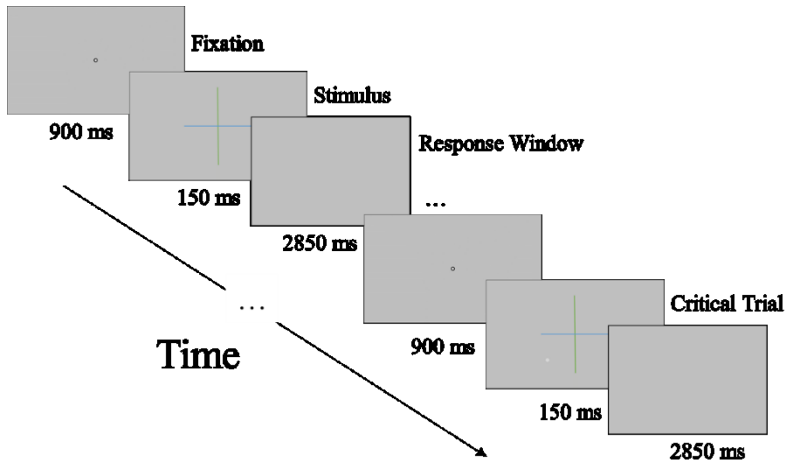

Figure 19. Visual task-visual critical signal task design.

Auditory Task-Auditory Critical Stimulus

The experimental task included the use of a rapid serial auditory presentation (RSAP) task, similar to that used by (Murphy, Fraenkel, & Dalton, 2013). Specifically, letters were presented rapidly to participants from speakers located offset to the left and right of the main computer, equidistant from the participant. Participants were asked to identify whether one of two specific targets were present during the task. Similar to the paradigm used by Murphy et al., targets included the letters "T" and "P". Each stimulus was presented for 240 ms, followed by 10 ms of silence such that each presentation lasted a total of 250 ms. Trials consisted of 6 spoken letters and always contained one of the two target letters. Only the high load condition described by Murphy et al was used, where subjects heard all different letters and responded to whether or not the target letter included was a T or a P (Figure 20). The auditory CS (a shape word) was presented about 80% of the way through the task. The auditory CS was presented at an intensity of about 5 dB lower than the experimental task, and took the place of one of the RSAP letters coming from either the left or the right speaker.

Figure 20. Auditory task – auditory critical signal design.

Crossmodal Inattentional Insensitivity Tasks

Crossmodal tasks were identical to intramodal tasks however, crossmodal CSs were presented only in the opposite modality to the central task. All other methods were the same as for the intramodal tasks. The combination of two intramodal tasks, and two crossmodal tasks created four conditions or groups in which each group received a different task first. Specifically, groups were defined by the task and the modality of their first critical signal trial (pure inattentional insensitivity) resulting in the following four groups: 1) visual task with a visual critical signal, 2) auditory task with an auditory critical signal, 3) visual task with an auditory critical signal, and 4) auditory task with a visual critical signal.

Working Memory Capacity

Working Memory Capacity (WMC) was measured using the Operation Span (OSpan) task, presented via the Psychology Experiment Building Language program (PEBL: Mueller & Piper, 2014). In the OSpan task participants completed simple mathematical equations and verified if a presented solution was correct or not (for example: (10/2) + 1 = 8? and (10/2) + 1 = 6? show incorrect and correct operations respectively) followed by a to-be-recalled letter. Participants read the operation, responded to the solution and read the to-be-recalled letter (see Conway et al., 2005; Turner & Engle, 1989). Lists could include two, three, four, five, or six letters. Two scoring methods were employed: Partial Credit Load Scoring (PCLS) and Partial Credit Unit Scoring (PCUS). These methods are presented and compared in the results section.

Procedure

Participants in all four groups were presented both tasks (the visual task and the auditory task) but were instructed to only respond to the task specific to their condition for each run (as instructed by the researcher). CSs were only presented in the modality matching the participant's condition. Prior to the start of each condition, participants practiced the specific task associated with that condition and were provided feedback as to their performance, in the form of a green or red fixation circle. After practice, participants completed the experimental block. Participants were told that they would see and hear both tasks but that they should only respond to the target task matching their condition, Table 11 shows the exact instructions that were read by the experimenter during various sections of the study.

Table 11. Exact instructions given to participants during the experiment

Instruction Set	Exact Instructions
OSpan Instructions	For this part of the study, the experiment will walk you through everything you need to do. You'll be completing what's called an OSPAN task, which is a task that will ask you to verify simple math problems while holding a string of letters in your head. There will be practices for all of the sections and the screen will have instructions for everything. I'll be here the whole time so just let me know if you have questions.
Main Overview	During this task, you will be presented with two different components, a visual part that involves monitoring a fixation cross, and an auditory part that involves monitoring a stream of spoken letters. The fixation cross has two arms, one will be blue and one will be green but also one arm will be longer than the other one and the stream of letters will contain 6 letters and one of them will always be either a P or a T. When you are doing the cross task, you'll respond by identifying which arm (vertical or horizontal) is the longer one. When you are doing the auditory task, you'll be listening to the stream of letters and responding by identifying whether there was a P or a T present.
Auditory Task Prioritized	During this block, we are asking you to only respond to the **auditory stream** while you monitor the cross. As a reminder, your goal for the auditory task is to respond by pressing the letter P if the stream contains a P and pressing the letter T if the stream contains a T. We will have a short practice where you will be given feedback before the real task. Do you have any questions?
Visual Task Prioritized	During this block, we are asking you to only respond to the **visual stream** while you monitor the letters. As a reminder, your goal for the visual task is to respond by pressing the letter H if the horizontal arm is the longer arm and V if the vertical arm is the longer arm. We will have a short practice where you will be given feedback before the real task. Do you have any questions?

Participants received a total of 10 blocks, each with between 6 and 9 trials. The first block always consisted of 9 trials and the CS was always presented on the 7th trial. The subsequent blocks only contained CSs in 6 of the 8 blocks, divided attention blocks. On the final trial, the CS was also always present, and participants were instructed for the final block not to perform either task, but to only monitor in order to assess their ability to notice the CS when not immersed in the primary task. Immediately following each block, participants were asked how engaged they felt with the task (between 0 and 10 where 0 would indicate their not being engaged at all with the task and 10 would indicate that they were extremely engaged in the task). Participants were asked whether they had noticed anything strange in the final trial of the task. Participants who did not report seeing the CS were considered non-noticers. Participants who stated that they did notice something were considered noticers and were asked if they were able to describe the CS (either the shape or word said and the location), which was then noted by the researcher. After this interview, participants completed the 8 divided-attention blocks and the final, full attention, block. The same post-block interview method was employed following each block. After completing all 10 blocks, participants completed the remaining three task-critical stimulus conditions in separate runs. Figure 21 shows the experimental design flow for one example run.

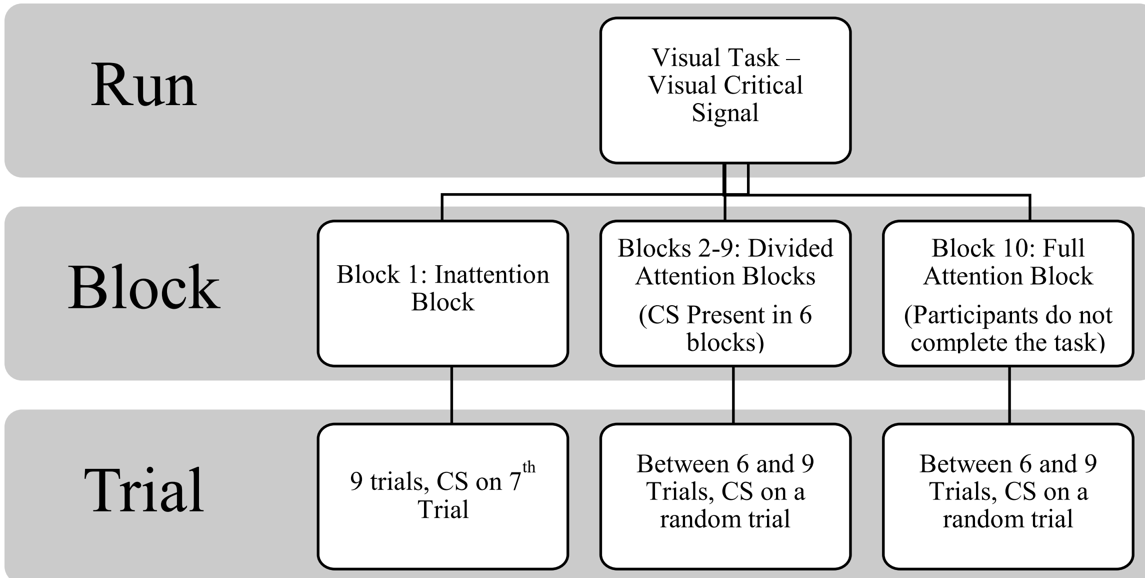

Figure 21. Experimental design using the Visual Task – Visual Critical Signal run as an example

2.3.4 Results

Operation Span Data Validation and Scoring

Prior to analysis of the experimental task, OSpan results were assessed for validity. Participants who scored less than 80% on the math portion of the OSpan were removed from analysis. Of the 155 participants who provided OSpan data, 18 participants were removed due to low performance on the math portion, leaving 137 with usable OSpan measures (some of whom may not have provided demographic information).

The OSpan task can be scored in multiple different ways. For the current experiment, we chose to use the partial-credit load scoring (PCLS) method described by Conway et al., (2005), where the score is the sum of all correctly recalled items divided by the total possible number of items as opposed to the partial-credit unit scoring (PCUS) method which computes the score as the average proportion of all correctly recalled items. Both methods were found by Conway et al. (2005) to be reliable across various span tasks and their derivations for one subject are shown in Table 12.

Table 12. Results for two different scoring procedures for one subject by trial with math score

Trial	Number of Items	Number of Correctly Recalled Items	Average Number of Correctly Recalled Items	Math Score
1	2	2	1	1
2	2	2	1	1
3	6	6	1	1
4	3	3	1	1
5	5	3	0.6	1
6	4	2	0.5	0.75
7	7	2	0.285714286	1
8	5	3	0.6	1
9	7	7	1	1
10	4	3	0.75	0.75
11	3	3	1	0.666667
12	6	6	1	1
13	7	7	1	1
14	6	5	0.833333333	1
15	3	3	1	1
16	4	1	0.25	1
17	5	2	0.4	1
Trials	Total Number of Items	Partial Credit Load Score (sum of all above items divided by total number of items)	Partial Credit Unit Score (sum of all above items divided by trials)	Math Score
17	79	0.759493671	0.777591036	0.950980412

Furthermore, within our current dataset the PCLS method resulted in a larger range of scores than the PCUS method (.73 as opposed to .67). In the studied population, both scoring methods resulted in negative skewness, but the PCLS method was slightly closer to normal than the PCUS method (-.742 as opposed to -.900).

Demographic Distribution by Groups

Of the 137 participants who provided usable OSpan data, 95 also completed the inattentional insensitivity portion of the task. **Error! Reference source not found.** shows the mean age, gender (where 1 denoted "Male" and 2 denotes "Female"), and working memory capacity score by first run condition with standard deviations shown in parentheses. A one-way analysis of variance indicated that there were no significant differences in any of these three metrics by group, *F* and *p* values are included in Table 13.

Table 13. Mean age, gender, and working memory capacity score by first run condition and type (SD values are shown in brackets)

Type	Condition	Usable N	Age	Gender	WMC Score
Intramodal Inattentional Insensitivity	Auditory Task – Auditory CS	25	19.32 (2.05)	1.84 (0.37)	0.75 (0.14)
	Visual Task – Visual CS	22	20.09 (2.73)	1.65 (0.49)	0.74 (0.16)
Crossmodal Inattentional Insensitivity	Auditory Task – Visual CS	25	20.20 (2.94)	1.60 (0.50)	0.71 (0.15)
	Visual Task – Auditory CS	23	19.50 (1.05)	1.85 (0.37)	0.77 (0.15)
Between groups analyses		***F*-Value**	0.806	1.996	0.825
		***p*-Value**	0.494	0.12	0.484

Overall Results for Critical Trial

Overall, across all conditions 78 subjects (82%) did not notice the critical signal on the first trial. Table 14 shows sensitivity by condition type (intramodal inattentional insensitivity or crossmodal inattentional insensitivity) for the first trial. Participants receiving CSs in the same modality as their main task were significantly more likely to notice the stimulus than those who received a CS in the opposite modality as their main task based on a Chi-Square analysis, $X^2(1, N = 90) = 4.44$, $p = .035$.

Table 14. Sensitivity by paradigm type and condition (Run 1 only)

Type	Condition	Insensitive	Sensitive
Intramodal Inattentional Insensitivity	Auditory Task – Auditory CS	68%	32%
	Visual Task – Visual CS	82%	17%
Crossmodal Inattentional Insensitivity	Auditory Task – Visual CS	92%	8%
	Visual Task – Auditory CS	86%	13%

A one-way analysis of variance (ANOVA) indicated that sensitivity for the first trial of each run significantly increased by run, *F(*3, 246) = 17;.48, p = .000. Post hoc comparisons using the Tukey HSD test indicated no significant differences (p = .06) between Run 1 (*M* = .179, *SD* = .385) and Run 2 (*M* = .337, *SD* = ..475), and no significant differences between Run 3 (*M* = .633, *SD* = .490) and Run 4 (*M* = .778, *SD* = .424), but significant differences between Runs 1 and 2, and Runs 3 and 4 as indicated in Figure 22. It should be noted here that, in almost all cases, the CS used in Run 2 was the opposite of the CS used in Run 1. Runs 3 and 4 would therefore have been the second time that a participant encountered the same CS which likely accounts for the significant jump in noticing after Run 2 and the lack of a significant difference from Run 1 to Run 2.

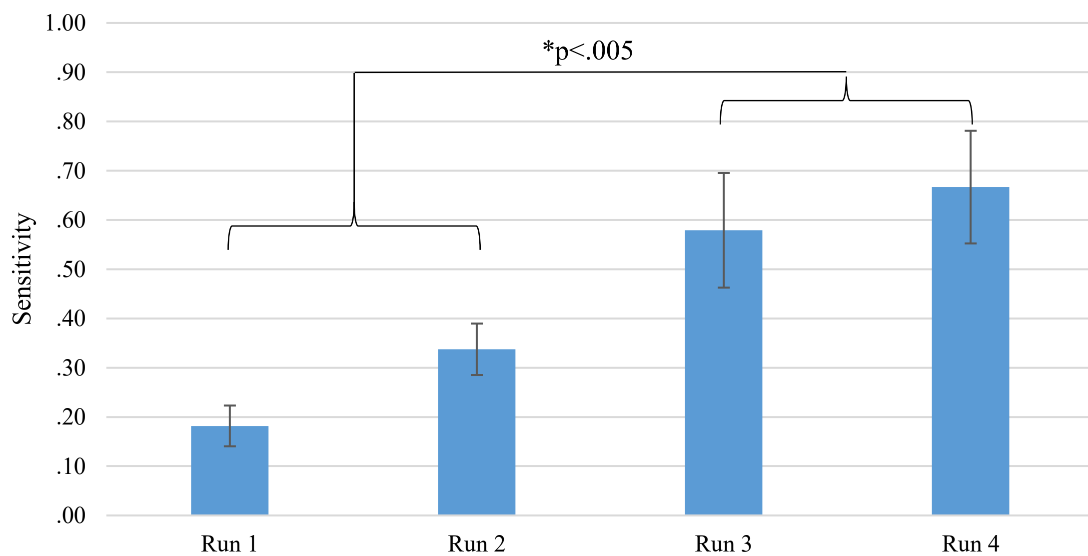

Figure 22. Sensitivity by run (where each run denotes one condition with ten blocks)

Working Memory Capacity

Results of curve estimation parameter estimates indicate no significant linear or quadratic effects of Operation Span (OSpan) on noticing in the first trial overall, $F(1,92) = .017$ $p = .895$, and $F(2,91) = .121$, $p = .449$, respectively. Further, there were no significant quadratic or linear effects of OSpan by condition.

As WMC as a continuous variable did not reach significance, participants were split into three equal groups based on all usable OSPANs. Low WMC participants were any with scores less than 68% and high WMC groups were any participants with scores higher than 84%. Table 15 shows the mean OSpan score for each category of WMC.

Table 15. Mean OSpan score for each categorical WMC level with standard deviation

WMC Level	Usable N	Mean	Std. Deviation
Low	30	0.56	0.10
Medium	28	0.76	0.05
High	36	0.89	0.04

A Chi-square analysis indicated no significant differences in noticing by WMC level, $X^2(2, N=94) = .713$, $p=.700$, as shown in Figure 23.

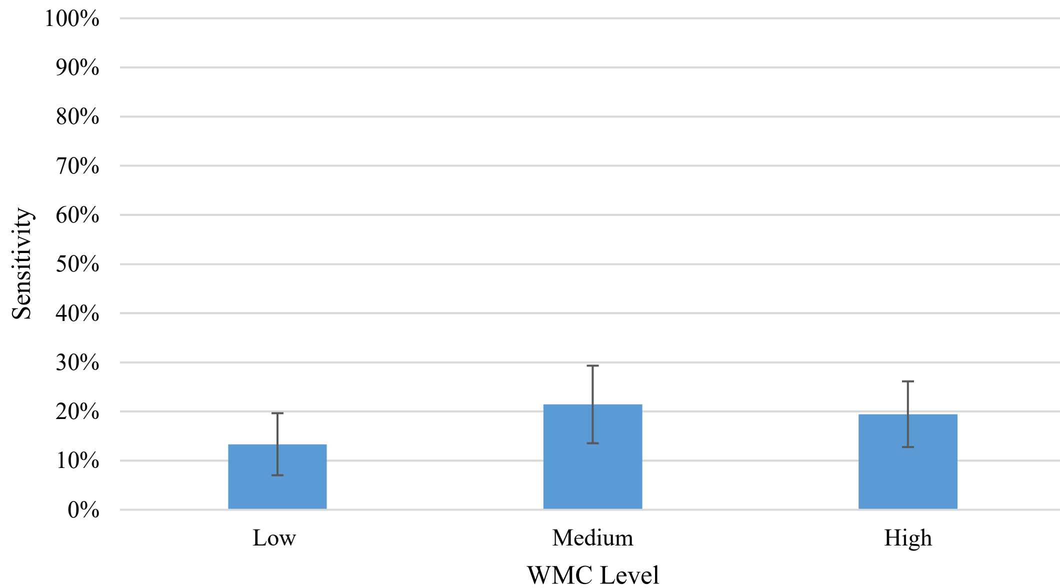

Figure 23. Sensitivity by WMC level

However, a Chi-Square analysis indicating significant differences in noticing by paradigm type (IIS vs CIS) was observed for those with High WMC, $X^2(1, N=34) = 3.848$, $p = .050$ (no significant effects were found for Low or Medium WMC participants). Figure 24 shows that, for those with High WMC, sensitivity was higher when the CS was presented in the same modality as the main task, as opposed to the opposite modality.

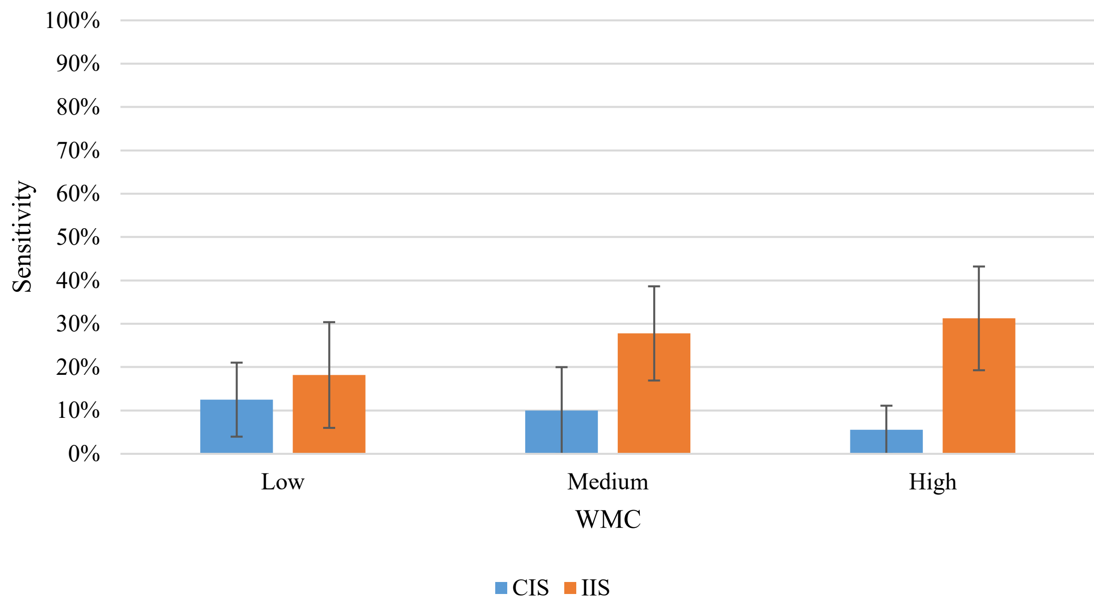

Figure 24. Sensitivity by working memory capacity by paradigm type (error bars represent standard error)

Furthermore, a univariate ANOVA of trials taken to notice the CS by WMC level (3: Low, Medium, High) and paradigm type (2: intramodal inattentional insensitivity [IIS], crossmodal inattentional insensitivity [CIS]) indicated a significant interaction effect such that Medium and High WMC participants took less trials to notice the CS in IIS paradigms than in CIS paradigms, $F(5,86) = 3.068$, $p = .014$, Figure 25 (trials taken to notice the CS could range from 1 meaning they noticed on the first trial, to 10 meaning that they didn't notice the CS until the last, full attention, trial). These two findings may indicate that higher WMC participants, seem to be more likely to filter out task/modality-irrelevant signals, and can more easily attend to irrelevant signals that are in the same modality as their primary perceptual task.

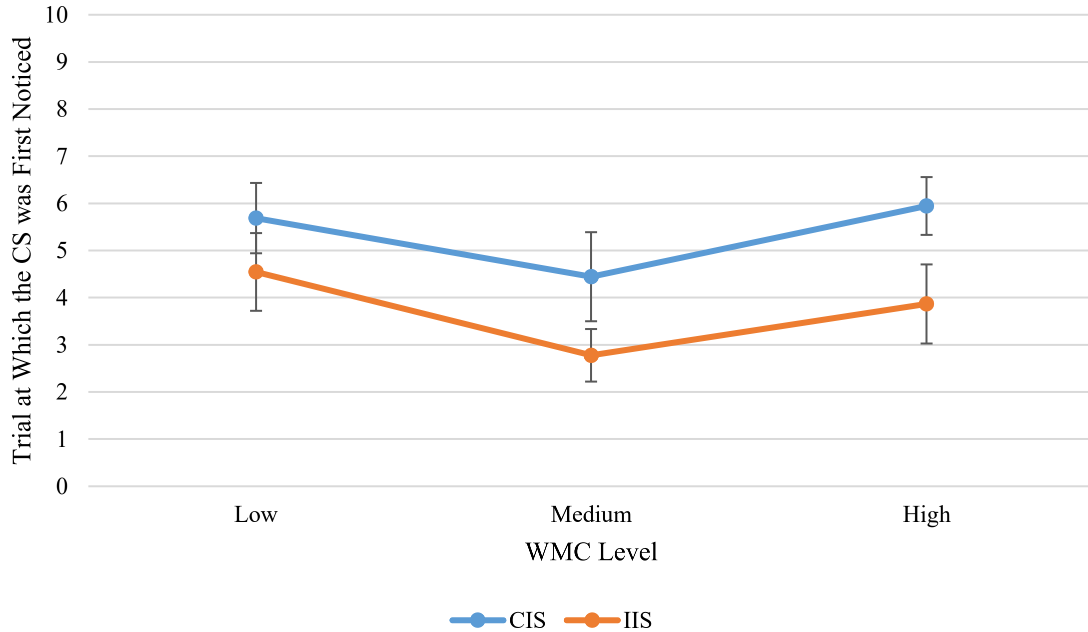

Figure 25. Number of trials before noticing by WMC by paradigm type (error bars represent standard error)

Additional Chi-Square analyses indicated no significant effects of critical signal type on noticing or the number of trials at which participants first noticed the CS.

Additional Predictors of Interest

In addition to condition and working memory capacity, experience with psychological terms such as "Inattentional Blindness", personality traits (as assessed by the Ten-Item Personality Index: Gosling, Rentfrow, & Swann, 2003), perceptions of difficulty on the primary task, and belief in ghosts (as a nod to the original work of

Cornell (1959) were also analyzed. Table 16 shows the predictors, outcomes and statistics for all predictors on noticing in the critical trial.

Table 16. Relationship between additional predictors of interest and inattentional insensitivity

Experience with Psychological Terms		
Predictor	**Outcome**	**Statistic**
Previous experience with the term "Inattentional Insensitivity"	No effect	Chi-Square Test: X^2 (2, N = 84) = 2.421, *p* = .298
Previous experience with the term "Inattentional Blindness"	No effect	Chi-Square Test: X^2 (2, N = 84) = 4.281, *p* = .118
Previous experience with the term "Inattentional Deafness"	Marginal Effect	Chi-Square Test: X^2 (2, N = 84) = 5.650, *p* = .059
Personality Traits as Assessed by the TIPI		
Predictor	**Outcome**	**Statistic**
Extraversion	No effect	Chi-Square Test: X^2 (12, N = 84) = 13.916, *p* = .306
Agreeableness	No effect	Chi-Square Test: X^2 (9, N = 84) = 7.972, *p* = .537
Conscientiousness	No effect	Chi-Square Test: X^2 (10, N = 84) = 9.383, *p* = .496
Emotional Stability	No effect	Chi-Square Test: X^2 (11, N = 84) = 14.344, *p* = .215
Openness	No effect	Chi-Square Test: X^2 (8, N = 84) = 4.395, *p* = .820
Difficulty of Primary Task		
Predictor	**Outcome**	**Statistic**
Perceived Difficulty of Auditory Task	No effect	Chi-Square Test: X^2 (10, N = 84) = 13.114, *p* = .217
Perceived Difficulty of Visual Task	No effect	Chi-Square Test: X^2 (10, N = 84) = 12.539, *p* = .185
Performance on Primary Task	No Effect	Chi-Square Test: X^2 (1, N = 76) = .151, *p* = .698
Reported Engagement	No Effect	Chi-Square Test: X^2 (7, N = 95) = 5.018, *p* =.658
Paranormal Belief		
Predictor	**Outcome**	**Statistic**
Belief in ghosts on noticing	No effect	Chi-Square Test: X^2 (1, N = 84) = .150, *p* = .699

No effect of previous experience with any psychological terms related to inattentional insensitivity (as a proxy for knowledge of the paradigm) were predictive of noticing, nor was the perceived difficulty of the main task. Analysis of the TIPI showed no significant effects of any measures. Finally, there was no effect of belief in ghosts on noticing.

2.3.5 Discussion

Results from this study indicate that inattentional insensitivity can be present in both intramodal and crossmodal paradigms and in individuals with varying levels of working memory capacity (WMC). Overall, around 82% of participants did not notice a critical signal (CS) the first time it was presented in the context of a perceptually difficult visual or auditory task. This number is well within the range typically reported in other studies, though is slightly higher than the average (50-60%, depending on the paradigm, see Table 1 for reference). However, some variation in the level of inattentional insensitivity was observed across paradigms. Specifically, the highest level of insensitivity was 92% in the Auditory Task – Visual CS condition and the lowest levels (68%) were in the Auditory Task – Auditory CS condition, with similar levels (82% - 86%) in the two Visual Task conditions. Additionally, as was expected, sensitivity increased by run, where participants were more likely to notice signals in subsequent runs particularly in the 3rd and 4th run where they had already been exposed to the specific CS regardless of primary task.

The main goal of this study was to examine the effect of WMC on inattentional insensitivity and to assess whether there were differences in inattentional insensitivity by compatibility of the CS and main task (intramodal or crossmodal). We were particularly interested in whether WMC as a continuous variable could support or refute our previously stated single- and dual-route hypotheses. To this end, we assessed the predictive power of WMC (OSpan score) as a continuous variable for inattentional insensitivity by run type with both linear and quadratic regressions. Although visual trends indicated that quadratic regression lines for noticing by OSpan may fit the data (though inversely in the case of the Visual Task – Auditory CS run), statistical analyses were non-significant. In order to better assess our hypotheses, and to compare our results to previous studies, we split our dataset into High, Medium, and Low WMC groups. We will now discuss the specific hypotheses that were examined in this experiment.

The first hypothesis was split into three parts, based on WMC level. Part one stated that individuals with very high WMC would show high levels of inattentional insensitivity in both crossmodal and intramodal tasks as they would actively inhibit task-irrelevant processing (H_1). H_1 had partial support. Participants in the High WMC group did show high levels of insensitivity (or low levels of sensitivity) overall, but a slight trend was observed such that High WMC participants had lower levels of insensitivity (and noticed in fewer trials) in intramodal runs than in crossmodal runs. Part two stated that individuals with very low WMC would show high levels of inattentional insensitivity in crossmodal and intramodal tasks as they would have limited spare resources with which to processes task-irrelevant stimuli (H_2). H_2 was partially supported. Participants in the Low WMC group showed high levels of insensitivity, and took somewhat longer to notice in both crossmodal and intramodal runs. The last part of the hypothesis stated that individuals with medium WMC would show the lowest levels of inattentional insensitivity across all tasks (H_3). H_3 was unsupported. Participants in the Medium WMC group did not show significantly different levels of inattentional insensitivity.

In addition to the main hypotheses tested, a second set of hypotheses was presented as alternates. The first part of the second hypothesis set stated that individuals with very high WMC would have the lowest levels of inattentional insensitivity as they would have spare resources to process additional information ($H_{1\text{-}0}$). $H_{1\text{-}0}$ was unsupported. High WMC participants did not have significantly different levels of inattentional insensitivity from other groups, nor could their levels of insensitivity be termed "low", at only 19% sensitivity. The second part stated that individuals with very low WMC would have low levels of inattentional insensitivity because they would be unable to inhibit task-irrelevant information ($H_{2\text{-}0}$). $H_{2\text{-}0}$ was unsupported. Only 13% of Low WMC participants noticed the CS on the first run. The third part stated that individuals in with all levels of WMC would show lower levels of inattentional insensitivity in crossmodal than in intramodal trials ($H_{3\text{-}0}$). $H_{3\text{-}0}$ was unsupported. Where there were differences in sensitivity, intramodal CSs were noticed significantly more often than were crossmodal CSs.

Finally, the null hypothesis stated that inattentional insensitivity would be present for a proportion of participants in all groups regardless of WMC or modality. The null hypothesis was partially supported. There were some differences in sensitivity by WMC and by CS type, but inattentional insensitivity was still present for a large proportion of all participants. In addition to the direct testing for the above hypotheses, other predictors of interest were assessed and are discussed in the following sections.

Previous Experience with Psychological Terms

Beanland & Pammer (2010) assessed previous knowledge of or familiarity with the term "Inattentional Blindness" and found that knowledge of the term or associated research did not necessarily predict noticing. Our results are in line with that finding and that of Bredemeier & Simons (2012), in that there were no significant effects of previous experience with the terms "Inattentional Insensitivity", or "Inattentional Blindness" on noticing in the first critical trial. Participants stating previous experience with "Inattentional Deafness" were slightly less likely to notice the CS on the first trial than those who said they had never heard of the term or were not sure (p = .059).

Personality Traits

Kreitz, Schnuerch, Gibbons, & Memmert (2015) found that most personality traits such as extraversion, neuroticism, agreeableness, motivation, schizotypy, absorption, behavioral inhibition, and behavioral activation were unrelated to inattentional blindness, but that openness (as measured by the Big Five Inventory) was negatively related to inattentional blindness. They concluded that "individuals that are open to new experiences in regard to interests, impressions, and ideas are also more "open" to unexpected objects" (Kreitz et al., 2015, p. 10). However, this did not hold true in our dataset. We found no significant differences in noticing by any personality traits. But, it is important to note here that we used a different inventory, the Ten Item Personality Index (TIPI).

Primary Task Performance and Perceptions of Difficulty

Simons & Jensen (2009) observed that although the demand of the primary task has been shown to affect inattentional blindness, the ability of participants to perform the task does not. Specifically, that tasks must reach a critical level of perceptual demand in order to create the appropriate paradigm to elicit some level of inattentional insensitivity. But that how well participants are able to perform the main task has no effect on sensitivity. Those findings are echoed in the current study. Both tasks were designed to be very perceptually demanding, but all participants practiced until they were able to perform reasonably well. Neither perceptions of the difficulty of the task (visual or auditory), reported engagement, nor performance on the primary task predicted inattentional insensitivity.

Paranormal Beliefs

Richards, Hellgren, & French (2014) investigated the relationship between absorption (susceptibility to highly focused attentional states), paranormal activity, and inattentional blindness, positing that the relationship between absorption and inattentional blindness, and absorption and paranormal beliefs may mediate the relationship between paranormal beliefs and inattentional blindness. The relationship was described by Richards et al. such that those who believed that they had experienced a paranormal event might actually have been displaying inattentional blindness, an example being the, "…apparently inexplicable movement of an inanimate object. Often a mundane explanation would plausibly solve the mystery by simply assuming that the claimant had failed to process some aspect of the original situation (e.g., the presence of an agent to move the object).” (Richards et al., 2014, p. 4). The current study did not support this hypothesis, as there was no relationship between reported belief in ghosts and inattentional insensitivity.

Limitations of the Current Study

It is important to note that, since this study did not include eye-tracking, it is not possible to ensure that all participants were looking at the screen at the time of the critical signal in this condition if their main task did not require them to focus on the screen. All participants were specifically told to ensure that they monitored the secondary task, and were reminded to watch and listen at all times, but no data is available on their specific eye positions which may explain the slightly higher than the average rate of inattentional blindness in the Auditory Task – Visual CS run. Participants in this study were recruited from the George Mason University (GMU) undergraduate research participation pool, meaning that in general, they represented only the younger portion of the general population. The average age of participants in this study was 20.32, with students’ ages ranging from 18 to 39. This meant that we were unable to assess the effect of age on inattentional insensitivity. This young population coupled with the fact that all participants in this study had, inherently, completed at least some advanced education (all were enrolled at least at the bachelor level at GMU), mean that the WMC scores for this sample were likely higher than the average for the general population. Additionally, as most subjects were enrolled in at least one psychology course at the time, the sample was more likely than average to have knowledge of the type of paradigm investigated here (though, reported familiarity was not statistically related to sensitivity). Furthermore, because of the nature of the design of the study, four between subject groups with WMC as a covariate, it is possible that the sample is somewhat underpowered. Although trends are apparent, statistical significance was not reached, despite the relatively large overall sample size.

Alternate Theories

To attempt to further understand our findings as they related to WMC, we returned to the literature. In 2016, Beanland & Chan published an article further attempting to tease out the relationship between WMC and inattentional insensitivity. They continued to find no direct relationship but did note in their review that, where a relationship has been found in the past, the association seems to only emerge when the sample population a) includes a wide age range, particularly of older adults; b) has little or no experience with the primary task; and c) when the CS is highly salient as opposed to just salient. In the case of the current study, the age range was restricted to a college population, participants had some practice on the primary task, and the CS was not of a particularly high salience. The suggestion here is that studies that find a direct relationship between WMC and inattentional insensitivity may be misattributing the perceptive or attentional abilities of their sample.

Parts of the hypotheses for the current study relied on multiple resource theory, believing that, if participants were overloaded in on modality they would be unable to process stimuli in that modality but able to process crossmodal signals. This was not the case in this sample. Rather, this sample points to a modality independent theory, such as Lavie's Load Theory (Lavie, Hirst, De Fockert, & Viding, 2004). Load Theory states that the extent to which irrelevant distractors are processed depends on whether perceptual capacity is reached, implying that individuals with higher perceptual capacity should show lower levels of inattentional insensitivity. This hypothesis is supported by findings from Swettenham et al (2014) and Tillmann & Swettenham (2017) related to children with autism spectrum disorder (ASD). Children with ASD are thought to have heightened perceptual processing capacities than the general population. In these two studies, Swettenham and colleagues showed that children with ASD had decreased inattentional insensitivity in intramodal and crossmodal paradigms.

With these findings in mind, we returned to theories inattentional insensitivity as related to task-induced attentional set. Various studies including Simons et al (2015), Koivisto & Revonsuo (2007), Most et al (2001), and Beanland & Chan (2016) have found that CS's are always more likely to be noticed when they are in the same attentional set as that primary task. That is, using the example of the gorilla task, participants who are told to count passes by players in white t-shirts are always more susceptible to inattentional blindness for the black gorilla than when they are told to count the passes by players in black t-shirts.   This was true for the current study, participants were significantly more likely to notice signals that were presented in the same modality (and therefore, attentional set) as the primary task. Furthermore, Kreitz, Furley, Memmert, & Simons (2016) hypothesized that:

> "… the effect of task-induced attention sets would be stronger for people with higher working memory capacity. Compared with participants with a lower working memory capacity, they should be more likely to notice unexpected

> objects that match the target items and less likely to notice those that differ from the target items." (p. 388)

This hypothesis exactly predicts the interaction effect found in the current dataset: that at high and medium levels of WMC participants are more likely to notice intramodal than crossmodal signals, but that this effect is not present at low WMC levels. These findings are also supported by research from Colflesh & Conway (2007) who investigated the controlled attention theory of working memory as it related to individual differences in divided attention during a dichotic listening task. The authors found that increases in WMC increased the ability to control focus of attention, or the spotlight of attention, such that high WMC participants could flexibly "zoom in or zoom out" depending on the difficulty and demands of the primary task.

2.3.6 Conclusion and Future Directions

Inattentional insensitivity is a phenomenon present for some proportion of the population across tasks, signal modalities and individual differences in working memory capacity. The exact interactions between these factors are difficult to assess due to the nature of inattentional insensitivity paradigms. They may be affected by personal factors such as previous experience with the paradigm or research, executive function or control, personality, or other state-related absorption, fatigue, or boredom. They may also be affected by task factors such as the inherent nature of the paradigm such that only the very first exposure can be truly considered an inattentional sensitivity measure. Further, the perceived or actual difficulty of the task, and the perceptual demands imposed by the task also impact inattentional insensitivity. The current findings indicate that there is a complex relationship between WMC and inattentional insensitivity which varies by task to CS modality. This relationship is best explained by the theory of controlled attention, such that task-induced attentional set predicts sensitivity, and does so more at higher WMC levels. Future research should attempt to further tease out this relationship by including a more robust population that varies more widely in working memory capacity, age, and education level. Future research should also consider including a wider battery of WMC tests and consider directly assessing attentional control. Researchers may also consider using physiological measures such as eye tracking and EEG or ERP analysis to determine whether inattentional insensitivity occurs at the perceptual or processing level across CS modalities.

3 GENERAL DISCUSSION

The purpose of the presented series of studies was to examine the design of in-vehicle alerting systems, the application of those systems, and to assess the reasons and circumstances in which these systems might fail operators. Two particular situations to which these failures apply are "Looked But Failed to See" (LBFTS) accidents, which make up around 6% of the 20% of all reported distraction-related accidents (according to the NHTSA FARS database), or failures of pilots to notice salient cockpit alarms. LBFTS accidents involve seemingly attentive drivers who fail to notice a salient signal, be it another car, a red light, or even a pedestrian, and can be considered a type of inattentional blindness. The case of pilot insensitivity involves pilots performing perceptually difficult, but highly trained, maneuvers during which they fail to notice a salient alarm indicating anything from failed landing gear to a proximity warning. These failures may happen at either the perceptual or the processing level due to high levels of perceptual (but not cognitive) load, and are not accounted for by ignoring, forgetting, or inability to notice under non-demanding load.

The first two studies presented in this series involved the design and application of in-vehicle alert systems with the intention of ameliorating insensitivity via effective and intuitive design. The first study used a form of psychophysical scaling called urgency coding to assess participants' perceptions of the urgency of varying parameters of signals in multiple modalities. We assessed visual, auditory, and tactile signals, varying color, flash rate, and word for visual signals, interpulse interval (IPI) for auditory and tactile signals, and frequency for auditory signals. We hypothesized that perceptions of urgency would increase with auditory frequency and with visual color, and would decrease with increases in flash rate and IPI. We further investigated relative signal utility, or the rate at which perceptions of annoyance increase as perceptions of urgency increase within parameters of each modality.

Results of the first study indicated that increases in perceptions of urgency were always associated with increases in annoyance but the degree to which they increased with each other did vary (their utility). Specifically, we found that changes in tactile IPI had a greater effect on urgency than on annoyance, relative to the same manipulations in the auditory IPI parameter indicating that tactile signals have a greater utility. Visual color showed very high utility, however it was not capable of producing perceptions of high urgency relative to auditory or tactile signals, making it a poor candidate for time-critical, highly urgent signals, but well-suited to non-critical changes related to the state of the vehicle, or an infotainment system. Tactile and auditory signals both resulted in similarly large ranges of urgency, and were capable of producing perceptions of high urgency, making them prime candidates for signals like forward collision warnings or lane deviation warnings. The higher utility of the tactile modality lends it well to systems that may produce high numbers of false alarms. However, there is a high possibility for tactile signals to be missed, a possibility that varies based on the implementation of the tactile warning system. For example, if the tactile warning system is in the seatbelt, there is the possibly that drivers who find the system annoying will stop using a seatbelt, or that those who do use the belt might be wearing large coats or bulky sweaters through which warnings might not be felt. Similarly, if the warning system is in base of the driver seat, drivers with thicker clothes or differing body types may not be able to feel warnings. Other candidates for the location of tactile warnings are the back of the driver seat, a location that may not always have contact with a driver's back, or in the steering wheel, where location of the driver's hands may affect perception of the warning. With these concerns in mind and following the current path of development for in-vehicle collision systems, we chose to further investigate auditory warnings in the context of a simulator-based hazard scenario in the second study.

The second study followed up on the results of the first study and from additional work in our labs designed to define key acoustic criteria for time-critical, highly urgent warnings. Specifically, we compared a forward collision warning (FCW) which met all five previously defined key criteria, an FCW we called GMU Prime, with an FCW which only met three of the key criteria, an FCW we called GMU Sub-Prime, and no FCW for a difficult, simulated, hazard scenario. Results of the second study indicated that the GMU Prime FCW increased the likelihood of an appropriate response to the collision scenario. The scenario was extremely difficult: a lead car switched lanes while both it and the participant vehicle were travelling around 65 mph to reveal a fully stopped vehicle. Although we did not see significant differences in collisions by FCW type or compared to no FCW, for those who did collide, the speed at which they collided was significantly lower than had they received no warning at all. We also saw some evidence that presenting a good alert, consistently, elicited better responses over time than did a poor alert or inconsistent alerts. Results of this study indicate that auditory FCWs can be designed using rigorous, psychophysically-based, testing methods, to be intuitive and effective for unanticipated, time-critical, highly urgent collision scenarios. But, that the collision must meet all five of the pre-defined key criteria: a peak-to-total time ratio of .7 or greater, at least three harmonic components, a base frequency of between 1000 Hz and 2500 Hz, and a pulse duration of 200 ms or less. Importantly we presented these alerts at an intensity that should be hard, if not impossible to miss, and during a perceptual task of intermediate difficulty, but some subjects still failed to respond appropriately. This may have been due to confusion, or an inability to re-orient to the driving scenario, or just a lack of experience avoiding collisions in the context of a simulated scenario. However, it is possible that some participants did not properly process the signal, or the situation, due to the load of the concurrent tasks. The third study used a basic paradigm to investigate the types of situations in which critical signals might be missed, and whether individual differences in working memory capacity (WMC) might explain differences in sensitivity.

The goal of the third study was to investigate whether the relationship between WMC and inattentional insensitivity in crossmodal and intramodal tasks could be explained by a single- or a dual-route model. If the relationship were explained by a single-route model, we hypothesized that as WMC increased, individuals would be less likely to show susceptibility to inattentional insensitivity as they would have spare resources with which to process task-irrelevant signals. If the relationship were explained by a dual-route model, we hypothesized that individuals with low or high WMC would show similarly high levels of susceptibility to inattentional sensitivity, with medium WMC individuals being the least susceptible to inattentional insensitivity. In the case of the dual-route model, we hypothesized that low WMC individuals would be more susceptible to inattentional insensitivity because they would be unable to spare any resources due to the highly demanding nature of the secondary task, but that high WMC individuals' susceptibility would be explained by overactive inhibition of irrelevant stimuli. In this case, we might expect to see differences in medium or high WMC individuals based on the type of task: intramodal or crossmodal.

We used two different tasks, presented simultaneously to participants, but prioritized based on the condition type. The visual task involved determining which arm of a presented cross was longer, the horizontal arm or the vertical arm. The auditory task involved determining which of two letters was present in a string of six, rapidly presented, letters, a T or a P (a rapid serial auditory presentation, RSAP task). Each task was presented in blocks of ten for each run, and participants were told to monitor one task, while responding to the other. In addition to the main task, visual or auditory critical signals (CSs) were presented in 8 of the 10 blocks of each run. Visual CSs were shapes located in one of the four quadrants of the cross, and Auditory CSs were shape words, embedded in the left or right RSAP stream. The combination of tasks and CSs created four unique conditions intramodal auditory, intramodal visual, crossmodal auditory, and crossmodal visual. Participants were split into four groups, each receiving the conditions in a different order, so that we could compare the very first block across run type, the critical block. At the end of each block, participants were interviewed to determine their sensitivity to CSs. All subsequent blocks, and those in the three subsequent runs, were considered divided attention blocks, as participants had some idea that they were supposed to be looking or listening for something that was not part of the main task.

Results of the third study indicated that individuals at any WMC level could show susceptibility to inattentional insensitivity and that insensitivity could be present for both intramodal and crossmodal paradigms, but that rates were significantly lower for intramodal paradigms. We found that over 80% of participants were susceptible to inattentional insensitivity, but that, particularly for those with medium or high (but not low) WMC, susceptibility was more likely in crossmodal paradigms than in intramodal paradigms. This finding lends support to models of attentional control such that individuals with higher working memory focus their attention to include only primary task-related objects or events while selectively ignoring objects or events that are outside of their attentional set.

Overall, results from the present series of studies indicate that further research is needed to better untangle the relationship between individual differences, signal design, and inattentional insensitivity-related accidents. Future research should identify the specific, underlying processes that contribute to inattentional insensitivity, be they WMC, attentional control, absorption, etc., in order to better design systems that are tailored to their users. Additionally, with the introduction of highly automated and autonomous vehicles, it will be important to ensure that signal design considers not only trait, but state-based factors such as distraction and fatigue. The advent of these systems has implied a future in which operators may be involved in demanding perceptual tasks such like video games during which they may be inattentionally insensitive to all but the most intrusive signals.

The current research offers some insight into the types of solutions that may be required in future signal design: the ability to create signals whose implied urgency can be recognized intuitively, facilitating appropriate responses (or in some cases, the deflection of responses); design recommendations for time-critical, highly urgent warnings; and some understanding of the relationship between individual differences and sensitivity to salient (but not intrusive) signals in multiple modalities. Specifically, that alert designers should consider multimodal, appropriately mapped signals to ensure that they both responded to correctly, and noticed in the first place.

4 GENERAL DISCUSSION

The purpose of the presented series of studies was to examine the design of in-vehicle alerting systems, the application of those systems, and to assess the reasons and circumstances in which these systems might fail operators. Two particular situations to which these failures apply are "Looked But Failed to See" (LBFTS) accidents, which make up around 6% of the 20% of all reported distraction-related accidents (according to the NHTSA FARS database), or failures of pilots to notice salient cockpit alarms. LBFTS accidents involve seemingly attentive drivers who fail to notice a salient signal, be it another car, a red light, or even a pedestrian, and can be considered a type of inattentional blindness. The case of pilot insensitivity involves pilots performing perceptually difficult, but highly trained, maneuvers during which they fail to notice a salient alarm indicating anything from failed landing gear to a proximity warning. These failures may happen at either the perceptual or the processing level due to high levels of perceptual (but not cognitive) load, and are not accounted for by ignoring, forgetting, or inability to notice under non-demanding load.

The first two studies presented in this series involved the design and application of in-vehicle alert systems with the intention of ameliorating insensitivity via effective and intuitive design. The first study used a form of psychophysical scaling called urgency coding to assess participants' perceptions of the urgency of varying parameters of signals in multiple modalities. We assessed visual, auditory, and tactile signals, varying color, flash rate, and word for visual signals, interpulse interval (IPI) for auditory and tactile signals, and frequency for auditory signals. We hypothesized that perceptions of urgency would increase with auditory frequency and with visual color, and would decrease with increases in flash rate and IPI. We further investigated relative signal utility, or the rate at which perceptions of annoyance increase as perceptions of urgency increase within parameters of each modality.

Results of the first study indicated that increases in perceptions of urgency were always associated with increases in annoyance but the degree to which they increased with each other did vary (their utility). Specifically, we found that changes in tactile IPI had a greater effect on urgency than on annoyance, relative to the same manipulations in the auditory IPI parameter indicating that tactile signals have a greater utility. Visual color showed very high utility, however it was not capable of producing perceptions of high urgency relative to auditory or tactile signals, making it a poor candidate for time-critical, highly urgent signals, but well-suited to non-critical changes related to the state of the vehicle, or an infotainment system. Tactile and auditory signals both resulted in similarly large ranges of urgency, and were capable of producing perceptions of high urgency, making them prime candidates for signals like forward collision warnings or lane deviation warnings. The higher utility of the tactile modality lends it well to systems that may produce high numbers of false alarms. However, there is a high possibility for tactile signals to be missed, a possibility that varies based on the implementation of the tactile warning system. For example, if the tactile warning system is in the seatbelt, there is the possibly that drivers who find the system annoying will stop using a seatbelt, or that those who do use the belt might be wearing large coats or bulky sweaters through which warnings might not be felt. Similarly, if the warning system is in base of the driver seat, drivers with thicker clothes or differing body types may not be able to feel warnings. Other candidates for the location of tactile warnings are the back of the driver seat, a location that may not always have contact with a driver's back, or in the steering wheel, where location of the driver's hands may affect perception of the warning. With these concerns in mind and following the current path of development for in-vehicle collision systems, we chose to further investigate auditory warnings in the context of a simulator-based hazard scenario in the second study.

The second study followed up on the results of the first study and from additional work in our labs designed to define key acoustic criteria for time-critical, highly urgent warnings. Specifically, we compared a forward collision warning (FCW) which met all five previously defined key criteria, an FCW we called GMU Prime, with an FCW which only met three of the key criteria, an FCW we called GMU Sub-Prime, and no FCW for a difficult, simulated, hazard scenario. Results of the second study indicated that the GMU Prime FCW increased the likelihood of an appropriate response to the collision scenario. The scenario was extremely difficult: a lead car switched lanes while both it and the participant vehicle were travelling around 65 mph to reveal a fully stopped vehicle. Although we did not see significant differences in collisions by FCW type or compared to no FCW, for those who did collide, the speed at which they collided was significantly lower than had they received no warning at all. We also saw some evidence that presenting a good alert, consistently, elicited better responses over time than did a poor alert or inconsistent alerts. Results of this study indicate that auditory FCWs can be designed using rigorous, psychophysically-based, testing methods, to be intuitive and effective for unanticipated, time-critical, highly urgent collision scenarios. But, that the collision must meet all five of the pre-defined key criteria: a peak-to-total time ratio of .7 or greater, at least three harmonic components, a base frequency of between 1000 Hz and 2500 Hz, and a pulse duration of 200 ms or less. Importantly we presented these alerts at an intensity that should be hard, if not impossible to miss, and during a perceptual task of intermediate difficulty, but some subjects still failed to respond appropriately. This may have been due to confusion, or to a lack of trust in the automated system, or an inability to re-orient to the driving scenario, or just a lack of experience avoiding collisions in the context of a simulated scenario. However, it is possible that some participants did not properly process the signal, or the situation, due to the load of the concurrent tasks. The third study used a basic paradigm to investigate the types of situations in which critical signals might be missed, and whether individual differences in working memory capacity (WMC) might explain differences in sensitivity.

The goal of the third study was to investigate whether the relationship between WMC and inattentional insensitivity in crossmodal and intramodal tasks could be explained by a single- or a dual-route model. If the relationship were explained by a single-route model, we hypothesized that as WMC increased, individuals would be less likely to show susceptibility to inattentional insensitivity as they would have spare resources with which to process task-irrelevant signals. If the relationship were explained by a dual-route model, we hypothesized that individuals with low or high WMC would show similarly high levels of susceptibility to inattentional sensitivity, with medium WMC individuals being the least susceptible to inattentional insensitivity. In the case of the dual-route model, we hypothesized that low WMC individuals would be more susceptible to inattentional insensitivity because they would be unable to spare any resources due to the highly demanding nature of the secondary task, but that high WMC individuals' susceptibility would be explained by overactive inhibition of irrelevant stimuli. In this case, we might expect to see differences in medium or high WMC individuals based on the type of task: intramodal or crossmodal.

We used two different tasks, presented simultaneously to participants, but prioritized based on the condition type. The visual task involved determining which arm of a presented cross was longer, the horizontal arm or the vertical arm. The auditory task involved determining which of two letters was present in a string of six, rapidly presented, letters, a T or a P (a rapid serial auditory presentation, RSAP task). Each task was presented in blocks of ten for each run, and participants were told to monitor one task, while responding to the other. In addition to the main task, visual or auditory critical signals (CSs) were presented in 8 of the 10 blocks of each run. Visual CSs were shapes located in one of the four quadrants of the cross, and Auditory CSs were shape words, embedded in the left or right RSAP stream. The combination of tasks and CSs created four unique conditions intramodal auditory, intramodal visual, crossmodal auditory, and crossmodal visual. Participants were split into four groups, each receiving the conditions in a different order, so that we could compare the very first block across run type, the critical block. At the end of each block, participants were interviewed to determine their sensitivity to CSs. All subsequent blocks, and those in the three subsequent runs, were considered divided attention blocks, as participants had some idea that they were supposed to be looking or listening for something that was not part of the main task.

Results of the third study indicated that individuals at any WMC level could show susceptibility to inattentional insensitivity and that insensitivity could be present for both intramodal and crossmodal paradigms, but that rates were significantly lower for intramodal paradigms. We found that over 80% of participants were susceptible to inattentional insensitivity, but that, particularly for those with medium or high (but not low) WMC, susceptibility was more likely in crossmodal paradigms than in intramodal paradigms. This finding lends some support to the hypothesized dual-route model: that low WMC participants were unable to process additional signals due to the taxing nature of the main perceptual task, but that medium and high WMC individuals may have been inhibiting attention to task-irrelevant (crossmodal) signals.

Overall, results from the present series of studies indicate that further research is needed to better untangle the relationship between individual differences, signal design, and inattentional insensitivity-related accidents. Future research should identify the specific, underlying processes that contribute to inattentional insensitivity, be they WMC, attentional control, absorption, etc., in order to better design systems that are tailored to their users. Additionally, with the introduction of highly automated and autonomous vehicles, it will be important to ensure that signal design considers not only trait, but state-based factors such as distraction and fatigue. The advent of these systems has implied a future in which operators may be involved in demanding perceptual tasks such like video games during which they may be inattentionally insensitive to all but the most intrusive signals.

The current research offers some insight into the types of solutions that may be required in future signal design: the ability to create signals whose implied urgency can be recognized intuitively, facilitating appropriate responses (or in some cases, the deflection of responses); design recommendations for time-critical, highly urgent warnings; and some understanding of the relationship between individual differences and sensitivity to salient (but not intrusive) signals in multiple modalities. Specifically, that alert designers should consider multimodal, appropriately mapped signals to ensure that they both responded to correctly, and noticed in the first place.

REFERENCES

Arndt, S. R., Wood, C. T., Delahunt, P. B., Krauss, D., & Wall, C. T. (2006). Who's in the Back Seat? A Study of Driver Inattention (Vol. 50, pp. 851–854). Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Sage Publications.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation*, *8*, 47–89.

Baldwin, C. L., & Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and older adults. *Psychology and Aging*, *26*(1), 85–91. http://doi.org/10.1037/a0020360

Baldwin, C. L., Eisert, J. L., Garcia, A., Lewis, B., Pratt, S. M., & Gonzalez, C. (2012). Multimodal urgency coding: auditory, visual, and tactile parameters and their impact on perceived urgency. *Work: A Journal of Prevention, Assessment and Rehabilitation*, *41*(0), 3586–3591.

Baldwin, C. L., & Lewis, B. A. (2014). Perceived urgency mapping across modalities within a driving context. *Applied Ergonomics*, *45*(5), 1270–1277.

Baldwin, C. L., & May, J. F. (2010). Loudness interacts with semantics in auditory warnings to impact rear-end collisions. *Transportation Research Part F*, *14*, 36–42.

Beanland, V., Allen, R. A., & Pammer, K. (2011). Attending to music decreases inattentional blindness. *Consciousness and Cognition*, *20*(4), 1282–1292.

Beanland, V., & Pammer, K. (2010). Gorilla watching: Effects of exposure and expectations on inattentional blindness (pp. 12–20). Presented at the ASCS09: Proceedings of the 9th Conference of the Australasian Society for Cognitive Science, Macquarie Centre for Cognitive Science Sydney.

Becklen, R., & Cervone, D. (1983). Selective looking and the noticing of unexpected events. *Memory & Cognition*, *11*(6), 601–608.

Bredemeier, K., & Simons, D. J. (2012). Working memory and inattentional blindness. *Psychonomic Bulletin & Review*, *19*(2), 239–244.

Brown-Iannuzzi, J. L., Hoffman, K. M., Payne, B. K., & Trawalter, S. (2014). The invisible man: Interpersonal goals moderate inattentional blindness to African Americans. *Journal of Experimental Psychology: General*, *143*(1), 33.

Chabris, C. F., Weinberger, A., Fontaine, M., & Simons, D. J. (2011). You do not talk about Fight Club if you do not notice Fight Club: Inattentional blindness for a simulated real-world assault. *I-Perception*, *2*(2), 150.

Clifasefi, S. L., Takarangi, M. K., & Bergman, J. S. (2006a). Blind drunk: The effects of alcohol on inattentional blindness. *Applied Cognitive Psychology*, *20*(5), 697–704.

Clifasefi, S. L., Takarangi, M. K., & Bergman, J. S. (2006b). Blind drunk: The effects of alcohol on inattentional blindness. *Applied Cognitive Psychology*, *20*(5), 697–704.

Conway, A. A., Kane, M., Bunting, M., Hambrick, D. Z., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. http://doi.org/10.3758/BF03196772

Cornell, A. (1959). An experiment in apparitional observation and findings. *Journal of the Society for Psychical Research*, *40*(701), 120–124.

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*, *124*(3), 367–372.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Dattel, A. R., Battle, A. N., Kossuth, L., Bifano, M., Decker, J., Majdic, K., … Sheehan, C. C. (2013). The Relationship between Inattentional Insensitivity of Visual, Tactile, and Olfactory Stimuli. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 57, pp. 853–857). SAGE Publications.

Dehais, F., Causse, M., Régis, N., Menant, E., Labedan, P., Vachon, F., & Tremblay, S. (2012). Missing critical auditory alarms in aeronautics: evidence for inattentional deafness? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, pp. 1639–1643). Sage Publications.

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to Detect Critical Auditory Alerts in the Cockpit Evidence for

Inattentional Deafness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(4), 631–644.

Drew, T., Võ, M. L.-H., & Wolfe, J. M. (2013). The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. *Psychological Science*, *24*(9), 1848–1853. http://doi.org/10.1177/0956797613479386

Edworthy, J., Hellier, E., Walters, K., Clift-Mathews, W., & Crowther, M. (2003). Acoustic, semantic and phonetic influences in spoken warning signal words. *Applied Cognitive Psychology*, *17*(8), 915–933.

Ferris, T. K., & Sarter, N. B. (2008). Cross-Modal Links Among Vision, Audition, and Touch in Complex Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(1), 17–26. http://doi.org/10.1518/001872008X250566

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030.

Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 179–185. http://doi.org/10.1016/j.trf.2008.11.002

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Graham, E. R., & Burke, D. M. (2011). Aging increases inattentional blindness to the gorilla in our midst. *Psychology and Aging*, *26*(1), 162.

Hannon, E. M., & Richards, A. (2010). Is inattentional blindness related to individual differences in visual working memory capacity or executive control functioning? Perception, 39(3), 309.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and new view. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 22, pp. 193–226). New York: Academic Press.

Hellier, E., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics*, *30*(2), 167–171. http://doi.org/10.1016/S0003-6870(97)00013-6

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors*, *35*(4), 693–706.

Hyman, I. E., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, *24*(5), 597–607.

Jang, P.-S. (2007). Designing acoustic and non-acoustic parameters of synthesized speech warnings to control perceived urgency. *International Journal of Industrial Ergonomics*, *37*(3), 213–223. http://doi.org/10.1016/j.ergon.2006.10.018

Kahneman, D. (1973). *Attention and effort*. Citeseer.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671.

Keidser, G., Ronnberg, J., Hygge, S., & Rudner, M. (2014). The Effect of Functional Hearing Loss and Age on Long-and Short-term Visuospatial 1 Memory: Evidence from the UK Biobank Resource.

Kennedy, K. D., & Bliss, J. P. (2013). Inattentional Blindness in a Simulated Driving Task (Vol. 57, pp. 1899–1903). Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications.

Koivisto, M., & Revonsuo, A. (2008). The role of unattended distractors in sustained inattentional blindness. *Psychological Research*, *72*(1), 39–48.

Kolers, P. (1969). Voluntary attention switching between foresight and hindsight. *Quarterly Progress Reports. Research Laboratory of Electronics, MIT*, (92), 381–385.

Kreitz, C., Schnuerch, R., Gibbons, H., & Memmert, D. (2015). Some see it, some don't: Exploring the relation between inattentional blindness and personality factors. *PloS One*, *10*(5), e0128158.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. http://doi.org/10.3758/BF03213897

Lewis, B. A., & Baldwin, C. L. (2012). Equating Perceived Urgency Across Auditory, Visual, and Tactile Signals. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1307–1311.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, *9*(3), 339–355. http://doi.org/10.1037/0882-7974.9.3.339

Lindenberger, U., & Mayr, U. (2013). Cognitive aging: is there a dark side to environmental support? *Trends in Cognitive Sciences*.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view." *Inhibition in Cognition*, *17*, 145–162.

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73*(6), 1780–1789.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual Attention*, *8*, 55–76.

Memmert, D. (2006). The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*, *15*(3), 620–627. http://doi.org/10.1016/j.concog.2006.01.001

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science*, *12*(1), 9-17.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250–259.

Murphy, D. R., Craik, F. I., Li, K. Z., & Schneider, B. A. (2000). Comparing the effects of aging and background noise of short-term memory performance. *Psychology and Aging*, *15*(2), 323.

Murphy, S., Fraenkel, N., & Dalton, P. (2013). Perceptual load does not modulate auditory distractor processing. *Cognition*, *129*(2), 345–355. http://doi.org/10.1016/j.cognition.2013.07.014

Nagayama, Y. (1978). Role of visual perception in driving. *IATSS Research*, *2*(1), 64–73.

Neisser, U. (1979). The control of information pickup in selective looking. *Perception and Its Development: A Tribute to Eleanor J. Gibson*, 201–219.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*(4), 480–494.

New, J. J., & German, T. C. (2014). Spiders at the Cocktail Party: An Ancestral Threat that Surmounts Inattentional Blindness. *Evolution and Human Behavior*.

O'Shea, D. M., & Fieo, R. A. (2014). Individual differences in fluid intelligence predicts inattentional blindness in a sample of older adults: a preliminary study. *Psychological Research*, 1–9.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*, 173.

Patterson, R., & Mayfield, T. (1990). Auditory warning sounds in the work environment [and Discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 485–492.

Remington, A., Cartwright-Finch, U., & Lavie, N. (2014). I can see clearly now: the effects of age and perceptual load on inattentional blindness. *Frontiers in Human Neuroscience*, *8*.

Richards, A., Hannon, E. M., & Derakshan, N. (2010). Predicting and manipulating the incidence of inattentional blindness. *Psychological Research*, *74*(6), 513–523.

Richards, A., Hellgren, M. G., & French, C. C. (2014). Inattentional blindness, absorption, working memory capacity, and paranormal belief. *Psychology of Consciousness: Theory, Research, and Practice*, *1*(1), 60–69. https://doi.org/10.1037/css0000003

Scholl, B. J., Noles, N. S., Pasheva, V., & Sussman, R. (2003). Talking on a cellular telephone dramatically increases "sustained inattentional blindness." *Journal of Vision*, *3*(9), 156–156. http://doi.org/10.1167/3.9.156

Seegmiller, J. K., Watson, J. M., & Strayer, D. L. (2011a). Individual differences in susceptibility to inattentional blindness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 785–791. http://doi.org/10.1037/a0022474

Seegmiller, J. K., Watson, J. M., & Strayer, D. L. (2011b). Individual differences in susceptibility to inattentional blindness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 785–791. http://doi.org/10.1037/a0022474

Simons, D. J. (2007). Inattentional blindness. *Scholarpedia*, *2*(5), 3244.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, *28*(9), 1059–1074.

Simons, D. J., & Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*, *16*(2), 398–403.

Slawinski, E. B., & MacNeil, J. F. (2002). AGE, MUSIC, AND DRIVING PERFORMANCE. *Psychomusicology: Music, Mind & Brain*, *18*(1), 123–131.

Stevens, S. (1971). Issues in psychophysical measurement. *Psychological Review*, *78*(5), 426.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*(3), 153–181. http://doi.org/10.1037/h0046162

Strayer, D. L., Cooper, J. M., & Drews, F. A. (2004). What do drivers fail to see when conversing on a cell phone? (Vol. 48, pp. 2213–2217). Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9*(1), 23–32. http://doi.org/10.1037/1076-898X.9.1.23

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Stevens' law. *The American Journal of Psychology*, 3–27.

Treat, J. R. (1980). A study of precrash factors involved in traffic accidents. *HSRI Research Review*.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635–654.

Valentijn, S. A. M., Van Boxtel, M. P. J., Van Hooren, S. A. H., Bosma, H., Beckers, H. J. M., Ponds, R. W. H. M., & Jolles, J. (2005). Change in Sensory Functioning Predicts Change in Cognitive Functioning: Results from a 6-Year Follow-Up in the Maastricht Aging Study. *Journal of the American Geriatrics Society*, *53*(3), 374–380. http://doi.org/10.1111/j.1532-5415.2005.53152.x

White, C. B., & Caird, J. K. (2010). The blind date: The effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident Analysis & Prevention*, *42*(6), 1822–1830.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* . (Third Edition). Prentice Hall.

Wolfe, J. M. (1999). Inattentional Amnesia Jeremy M. Wolfe. *Fleeting Memories: Cognition of Brief Visual Stimuli*, 71.

Zacks, R. T. (1989). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, *22*(22), 193–225.

BIOGRAPHY

Bridget A. Lewis received her B.S. in psychology from Virginia Polytechnic Institute and State University and her M.A. from George Mason University. Her research interests include multimodal display design and evaluation, the design of transportation systems for accessibility, designing for older adults, human factors in healthcare simulation, and data visualization for display design.