

In-vehicle Multimodal Interaction
an Approach to Mitigate Driver Distraction

by

Richa Mittal

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2015 by the
Graduate Supervisory Committee:

Ashraf Gaffar, Chair
John Femiani
Robert Gray

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

Despite the various driver assistance systems and electronics, the threat to life of driver, passengers and other people on the road still persists. With the growth in technology, the use of in-vehicle devices with plethora of buttons and features is increasing resulting in increased distraction. Recently, speech recognition has emerged as an alternative to distraction and has the potential to be beneficial. However, considering the fact that automotive environment is dynamic and noisy in nature, distraction may not arise from the manual interaction, but due to the cognitive load. Hence, speech recognition certainly cannot be a reliable mode of communication.

The thesis is focused on proposing a simultaneous multimodal approach for designing interface between driver and vehicle with a goal to enable the driver to be more attentive to the driving tasks and spend less time fiddling with distractive tasks. By analyzing the human-human multimodal interaction techniques, new modes have been identified and experimented, especially suitable for the automotive context. The identified modes are touch, speech, graphics, voice-tip and text-tip. The multiple modes are intended to work collectively to make the interaction more intuitive and natural. In order to obtain a minimalist user-centered design for the center stack, various design principles such as 80/20 rule, contour bias, affordance, flexibility-usability trade-off etc. have been implemented on the prototypes. The prototype was developed using the Dragon software development kit on android platform for speech recognition.

In the present study, the driver behavior was investigated in an experiment conducted on the DriveSafety driving simulator DS-600s. Twelve volunteers drove the simulator under two conditions: (1) accessing the center stack applications using touch only and (2) accessing the applications using speech with offered text-tip. The duration for which user looked away from the road (eyes-off-road) was measured manually for each scenario. Comparison of results proved that eyes-off-road time is less for the second scenario. The minimalist design with 8-10 icons per screen proved to be effective as all the readings were within the driver distraction recommendations (eyes-off-road time < 2sec per screen) defined by NHTSA.

ACKNOWLEDGMENTS

I am very grateful to Dr. Ashraf Gaffar for giving me the opportunity to work on this project and providing me the valuable guidance and continuous support. I am grateful to Dr. Robert Gray for kind courtesy extended to me by providing the access to Simulator lab to conduct the experiments. Also I would like to thank Dr. John Femiani for readily accepting to be my thesis committee member. I would like to thank my colleagues Tanvi Jahagirdar, Paresh Nakrani and all the members of EcoCar3 Innovation Team for constantly motivating and helping me out through the project.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
LIST OF FIGURES.....	vi
CHAPTER	
1 INTRODUCTION	1
2 BACKGROUND	4
What is Driver Distraction?.....	4
Sources of Driver Distraction.....	5
Distraction-Affected Crashes.....	7
Driver Behavior Detection and Support	12
US. DOT Recommendations.....	15
3 RESEARCH APPROACH	16
Today's Car: Computer On-Wheels.....	18
Advance Driver Assistive System	21
4 AUTONOMOUS VEHICLES	27
Autopilot	28
Drones.....	29
Subway	32
Autonomous Road Vehicle- Google Car	33
5 MOTIVATION FOR IMMI	40
Speech Recognition	40
Is Technology A Curse Or A Blessing?	43
6 MULTIMODAL INTERACTION	45
Human-Human Multimodal Interaction	45
Human-Computer Multimodal Interaction	46
7 IN-VEHICLE MULTIMODAL INTERACTION- OUR APPROACH	48

CHAPTER	PAGE
8 DESIGN	52
Navigation Model	52
User-Interface Design.....	56
9 RELATED PRIOR WORK	65
Experiment 1	65
Experiment 2.....	66
10 EXPERIMENT	70
Engineering-Based Metrics	70
Experiment 1	71
Experiment 2.....	73
11 THE OUTREACH EVENT	80
12 CONCLUSION	84
13 FUTURE WORK	85
REFERENCES.....	86
APPENDIX	
A GROWTH RATE OF CRASHES	90
B DRIVING SIMULATOR	92
C STATISTICAL ANALYSIS	95
D QUESTIONNAIRE FOR THE OUTREACH EVENT	104
E READINGS FOR EXPERIMENT 1	106
F READINGS FOR EXPERIMENT 2	108
G CONSENT FORM	110

LIST OF TABLES

Table		Page
1.	Police Reported Crashes and Crashes Involving Distraction, 2006 – 2010	7
2.	Fatal Crashes, Drivers in Fatal Crashes, and Fatalities, 2012	8
3.	Percentage Killed in Distraction-Affected Crashes, by Person Type, 2012.....	9
4.	Crash, Death, Vehicle, Population Growth Rate	74
5.	Experiment 2 Readings	87
6.	Experiment 1 Readings.....	88

LIST OF FIGURES

Figure		Page
1.	Crash, Death, Population Growth Rate & Motor Vehicle Rate- Year 1975 To 2008	1
2.	Crash Factors	2
3.	Percent Distribution of Drivers Involved in Fatal Crashes By Age	9
4.	Progressive Domain Analysis	16
5.	Trends In Muscle Car	18
6.	Computer-On-Wheels	18
7.	Autodesk Tool for Car Design	19
8.	Carmaker for Simulink for Model-Based Design, and Testing of Cars	19
9.	Nissan’s Car Manufacturing Plant	20
10.	Mercedes S Class	21
11.	Gartner’s Hype Cycle	29
12.	Gartner’s Hype Cycle for Human-Computer Interaction	30
13.	Car Models Made in the North America for U.S Market, 2014	32
14.	Human-Human Multimodal Communication	34
15.	Human-Computer Interaction	35
16.	Human-Vehicle Interaction	36
17.	Navigation Model: Screen to Screen	40
18.	Navigation Model: Multiple Facets	41
19.	Navigation Model: Superimposed Screens	41
20.	Navigation Model- Level 1	42
21.	Navigation Model- Level 2	43
22.	Navigation Model- Level 3	44
23.	Home Screen with Text	46
24.	Home Screen without Text	46
25.	Car Systems without Text	47

Figure	Page
26. Car Systems with Text	47
27. Media Player	47
28. Settings Screen.....	49
29. Air Conditioning System	50
30. GPS Screen Without Text.....	50
31. GPS Screen Without Text.....	50
32. Prototype Images	51
33. Phone Screen with and without Text	52
34. Driver's View of the Simulator with 10" Tablet Displaying 8 Icons.....	54
35. Small Screen with 24 Icons.....	56
36. Single Mode Vs Dual Mode	61
37. Route Overview.....	62
38. The Starting Point	62
39. The Left Turn.....	63
40. The Curved Road.....	63
41. Pedestrian Crossing.....	64
42. The Arrow Pointing Towards the Microphone used for Speech Recognition	65
43. Lane Deviation	66
44. Eyes Take-Off-Road Time: Touch Vs Speech With Text	68
45. The Outreach Event.....	69
46. Students Filling Out the Questionnaire at the Outreach Event.....	70
47. Students while Driving the Simulator	70
48. Driving Simulator.....	77
49. Driving Simulator Computer.....	77
50. Center Stack Screen Holder	78

CHAPTER 1

INTRODUCTION

Driving is the coordinated operation of mind and body for the movement of a vehicle. Today, driving being considered as an everyday activity for most of the people, still has an issue of safety. A study predicts that passenger-miles of travel will be faster than the growth of passenger miles in travel experienced during the 1990s, from 5 trillion miles in 2000 to 8.4 trillion in 2025 along with a corresponding rise in global travel [1]. Over the years, we have seen emerging technology for safer driving. Electronic stability control, collision avoidance systems, intelligent speed adaptation, and vehicle tracking systems can all help mitigate the threats to drivers [2]. Great improvements have been introduced to many aspects of modern cars, from better engines and chassis construction to higher vehicle stability, better wheels and tires, and better overall crash protection. Unfortunately, the total number of fatal crashes is still a problem despite the safety improvements in road and vehicle design. In 2011, there were nearly 34,400 transportation related deaths and over 2.2 million transportation-related injuries. Furthermore, in 2013, a total of 32,719 people died in 30,057 crashes involving 44,868 motor vehicles [3].

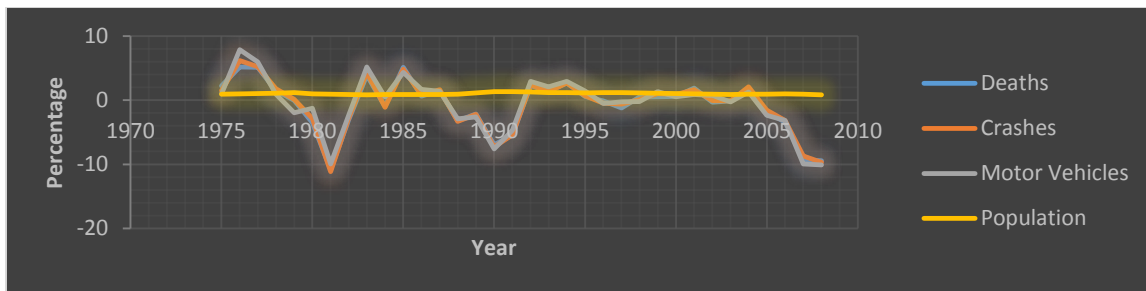


Figure 1: Crash Rate, Death rate, population growth rate & motor vehicle rate from year 1975 to 2008 [3], [4]

Figure 1 shows the relation between population growth rate, crash rate, death rate and vehicle rate from year 1975 to 2008 [3]. One could argue that the crash rate and death rate is proportional to rate of motor vehicles. Nonetheless, the problem seems to be more complex than just a correlation. However, despite the fact that transportation safety has improved resulting a significant decline in fatalities, yet keeping an average of 92 deaths per day in motor vehicle crashes in 2012, which is still a very huge loss [5].

One of the earliest driver distraction studies collected data between 1972 and 1975 and grouped into three “levels.” Level A was a collection of baseline data including vehicle registration and driver’s license information as well as surveys from the general population. Level B was a data set collected from police accident reports. A total of 2,258 crashes were investigated (crashes involving heavy vehicles and vehicles pulling trailers were not included). Level C data was an in-depth investigation of Level B data and included 420 crashes. For each crash in Level C, there was an investigation of human, environment, and vehicle factors that may have contributed to the crash.

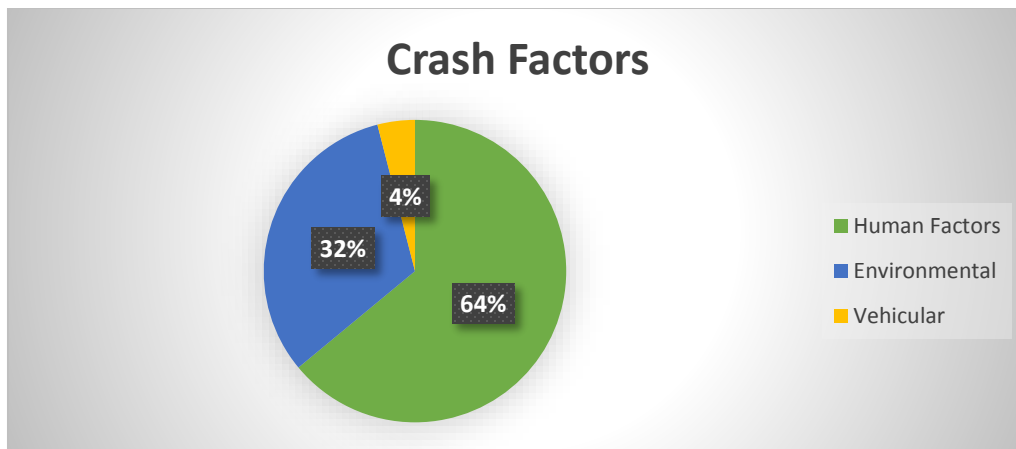


Figure 2: Crash Factors

As illustrated by Figure 2, the results of this study found that human factors were most often (71–93 percent of the analyzed cases) cited as the cause in the crashes, followed by environment (12–34 percent of the analyzed cases) and vehicle factors (5–13 percent of the analyzed cases). Five major categories of human direct causes were identified: recognition errors, decision errors, performance errors, critical non-performance errors, and non-accident/intentional involvement. Furthermore, five specific human causes were identified: improper lookout (18–23 percent), excessive speed (8–17 percent), inattention (10–15 percent), improper evasive action (5–13 percent), and internal distraction (6–9 percent). It can be seen that two of the five specific human causes were related to inattention and distraction [6].

Most of the crashes are due to driver's inattention. Four major categories of attentional impairments include alcohol, fatigue, age, and distraction. Alcohol contributes to approximately 40% of fatalities in US highway [7]. Fatigue is often found in the accidents involving young drivers and truck drivers because these drivers tend to adopt risky strategies to drive at night and/or lack good-quality sleep. Aging results in longer response time to hazards and more narrow field of attention in old drivers. Compared with the above three impairments, distraction, the fourth impairment, has become increasingly important with the introduction of in-vehicle technology (e.g., navigation systems, cell phones, and internet) and has drawn increasing attention from human factor researchers and policy makers in the area of transportation safety [8].

The study is majorly focused on analyzing the crashes due to driver distraction involving the use of any kind of device and identifying a mitigation approach. In the latter half of the study, multimodal interaction framework and techniques have been discussed. Chapter 2 provides more insight into driver distraction.

CHAPTER 2

BACKGROUND

What Is Driver Distraction?

As defined by the official US government website for distractive driving [9],

“Distracted driving is any activity that could divert a person's attention away from the primary task of driving. All distractions endanger driver, passenger, and bystander safety. These types of distractions include:

- Texting
- Using a cell phone or smartphone
- Eating and drinking
- Talking to passengers
- Grooming
- Reading, including maps-
- Using a navigation system
- Watching a video
- Adjusting a radio, CD player, or MP3 player”

As defined by National Highway Traffic Safety Administration [10],

“Distraction is anything that diverts the driver's attention from the primary tasks of navigating the vehicle and responding to critical events. To put it another way, a distraction is anything that takes your eyes off the road (visual distraction), your mind off the road (cognitive distraction), or your hands off the wheel (manual distraction).”

As defined by Governors Highway Safety Association,

“There are many distractions which may prevent a driver from focusing on the complex task of driving: changing the radio or a CD, talking to passengers, observing an event outside the vehicle, using an electronic device, etc. Navigational and other interactive devices, called telemat-

ics, in the vehicle are available in most vehicles and more will be available in the near future. These devices may also distract drivers.”

However, the American Automobile Association (AAA) Foundation challenges the notion that drivers are safe and attentive as long as their eyes are on the road and their hands are on the steering wheel. Mental distractions can dangerously affect drivers behind the wheel. Just because a driver’s eyes are on the road and hands are on the wheel does not mean that they are safe. Attention is key to safe driving, yet many technologies can cause drivers to lose focus of the road ahead. Hands-free and voice-command features, increasingly common in new vehicles, may create mental distractions that unintentionally provide motorists with a false sense of security about their safety behind the wheel [16].

Driver distraction may be characterized as any activity that takes a driver’s attention away from the task of driving. An examination of the crash data reveals that any distraction has the potential to cause or contribute to a crash. Thus, rolling down a window, adjusting a mirror, talking to other passengers, tuning a radio or dialing a cell phone are all contributing factors in crashes. Recent concerns about the potential safety implications of technology based distractions center on the magnitude and nature of demands.

Sources of Driver Distraction

Based on the literature review, distraction can be categorized as: visual, manual, auditory and cognitive. Visual distraction occurs when the driver focuses on some other task instead of looking at the road, in simple words “eye-off-road”. Auditory distraction occurs when the driver focuses their attention on auditory signals such as music rather than on the road environment, “ears-off-road”. Manual distraction occurs when drivers remove hands from the steering wheel to physically manipulate an object, “hands-off-road”. Cognitive distraction includes any thoughts that absorb the driver’s attention to the point where they are unable to focus on driving, “mind-off-road”. Any of the above type of distraction can lead to larger lane variation, more abrupt steering control, slower response to hazards and less efficient visual perception than attentive driving. Moreover, the four types of distraction can occur in combination and interact with each other.

However, cognitive distraction is the most difficult of the four sources of distraction to be assessed non-intrusively. As compared to the other activities studied (e.g., listening to the radio, conversing with passengers, etc.) we found that interacting with the speech-to-text system was the most cognitively distracting. This clearly suggests that the adoption of voice-based systems in the vehicle may have unintended consequences that adversely affect traffic safety [12]. In addition to this, research has proved that increased cognitive load leads to fewer saccades and less variations in lane position, but increasing the incidents of hard braking, failure to scan for potential hazards in the driving environment, failure to notice objects in the line of sight, and failures to stop at controlled intersections [13]–[15].

An enduring question concerns the ability of humans to multitask. As technological and informational capabilities of our environment increase, the number of available information increases, and hence the opportunities for complex multitasking increase. At the most basic level, we can say that this involves keeping track of the other entities on road, the steering wheel, the brake and accelerator. At a more complex level, this may involve shifting down to a lower gear while navigating a curve and talking with other passengers. There is a conscious trade off performing one task for the other, also performances in both tasks depend highly on their skill in the individual tasks [16]. Driving itself involves such complex tasks and operating in-vehicle devices is an add-on to it.

We know that cognitive load theory deals with learning and problem solving difficulty. Intrinsic cognitive load in contrast, is constant for a given area because it is a basic component of the material and is characterized in terms of element interactivity. The elements of most schemas must be learned simultaneously because they interact and it is the interaction that is critical. If, as in some areas, interactions between many elements must be learned, then intrinsic cognitive load will be high. In contrast, in different areas, if elements can be learned successively rather than simultaneously because they do not interact, intrinsic cognitive load will be low [17]. Lower the cognitive load, reduced is the distraction. Hence, we propose a user interface that is easy to operate without being required any special training.

Distraction-Affected Crashes

NHTSA crash data, the major components of inattention-related police reported crashes include:

- “Distraction” (attending to tasks other than driving, e.g., tuning the radio, speaking on a phone, looking at a billboard, etc.),
- “Looked but did not see” (e.g., situations where the driver may be lost in thought or was not fully attentive to the surrounds)
- Situations where the driver was drowsy or fell asleep.

All together, these crashes account for approximately 25 percent of police reported crashes. Distraction was most likely to be involved in rear-end collisions in which the lead vehicle was stopped and in single vehicle crashes. Crashes in which the driver “looked but did not see” occurred most often at intersections and in lane-changing/merging situations.

The most recent data available, 2010 data, show that 899,000 motor vehicle crashes involved a report of a distracted driver (17 percent of all police-reported crashes: fatal, injury only, and property-damage-only). As seen in Table 1, the percentage of all police-reported crashes that involve distraction has remained consistent over the past five years. On average, these distraction-related crashes lead to thousands of fatalities (3,092 fatalities or 9.4 percent of those killed in 2010) and injuries to over 400,000 people each year (approximately 17 percent of annual injuries).

Year	Number of Police Reported Crashes	Police-Reported Crashes Involving a Distracted Driver	Police-Reported Crashes Involving a Distracted Driver Using an Integrated Device	Police-Reported Crashes Involving a Distracted Driver Using Electronics
2006	5,964,000	1,019,000 (17%)	18,000 (2%)	24,000 (2%)
2007	6,016,000	1,001,000 (17%)	23,000 (2%)	48,000 (5%)
2008	5,801,000	967,000 (17%)	21,000 (2%)	48,000 (5%)
2009	5,498,000	957,000 (17%)	22,000 (2%)	46,000 (5%)
2010	5,409,000	899,000 (17%)	26,000 (3%)	47,000 (5%)

Table 1: Police Reported Crashes and Crashes Involving Distraction, 2006 – 2010 [18]

Of the 899,000 distraction-related crashes, 26,000 (3%) specifically stated that the driver was distracted while adjusting or using an integrated device/control. From a different viewpoint, of those 899,000 crashes, 47,000 (5%) specifically stated that the driver was distracted by a cell phone (no differentiation between portable and integrated cell phones). It should be noted that these two classifications are not mutually exclusive, as a driver distracted by the integrated device/control may have also been on the phone at the time of the crash and thus the crash may appear in both categories. While all electronic devices are of interest, the current coding of the crash data does not differentiate between electronic devices other than cell phones [18].

Today's 2011 National Occupant Protection Use Survey (NOPUS) shows that at any given daylight moment across America, approximately 660,000 drivers are using cell phones or manipulating electronic devices while driving, a number that has held steady since 2010. According to separate NHTSA data, more than 3,300 people were killed in 2011 and 387,000 were injured in crashes involving a distracted driver [19]. The number of people killed in distraction-affected crashes decreased slightly from 3,360 in 2011 to 3,328 in 2012. An estimated 421,000 people were injured in motor vehicle crashes involving a distracted driver, this was a nine percent increase from the estimated 387,000 people injured in 2011 [9]. Five seconds is the average time your eyes are off the road while texting. When traveling at 55mph, that's enough time to cover the length of a football field blindfolded [9]. A distraction-affected crash is any crash in which a driver was identified as distracted at the time of the crash. Federal estimates suggest that distraction contributes to 16% of all fatal crashes, leading to around 5,000 deaths every year [20]. In 2012, there were a total of 30,800 fatal crashes in the United States involving 45,337 drivers. As a result of those fatal crashes, 33,561 people were killed.

Total	Crashes	Drivers	Fatalities
	30,800	45,337	33,561
Distraction-Affected	3,050 (10% of total crashes)	3,119 (7% of total drivers)	3,328 (10% of total fatalities)

Table 2: Fatal Crashes, Drivers in Fatal Crashes, and Fatalities, 2012[21]

As per the information provided in Table 2, 3,050 fatal crashes occurred on U.S. roadways that involved distraction (10% of all fatal crashes). These crashes involved 3,119 distracted drivers, as some crashes involved more than one distracted driver. Distraction was reported for 7 percent (3,119 of 45,337) of the drivers involved in fatal crashes [21].

Another age group of distracted drivers to look at is the 20-to-29-year-old driver group. Drivers in their 20s make up 23 percent of drivers involved in all fatal crashes, 27 percent of the distracted drivers in fatal crashes and 34 percent of the distracted drivers that were using cell phones. Figure 3 illustrates the distribution of drivers by age for total drivers involved in fatal crashes, distracted drivers involved in fatal crashes, and distracted drivers on cell phones during fatal crashes [21].

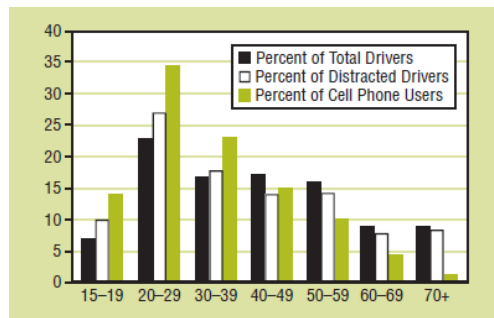


Figure 3: Percent Distribution of Drivers Involved in Fatal Crashes by Age, 2012[21]

In 2012, 84 percent of the fatalities in distraction-affected crashes involved motor vehicle occupants or motorcyclists. This compares to 83 percent of all motor vehicle crash fatalities involving occupants. Thus, the victims of distraction-affected crashes vary little from the victims of crashes overall.

Occupant			Non-Occupant			
Driver	Passenger	Total	Pedestrian	Cyclist	Other	Total
2,010 (60%)	778 (23%)	2,778 (84%)	434 (13%)	81 (2%)	25 (1 %)	540 (16%)

Table 3: Percentage Killed in Distraction-Affected Crashes, by Person Type, 2012[21]

Table 3 describes the role of the people killed in distraction-affected crashes in 2012. Distracted drivers were involved in the deaths of 540 non-occupants during 2012.

In September 2010, the NHTSA released a report on distracted driving fatalities for 2009. The NHTSA considers distracted driving to include the following distractions: other occupants in the car, eating, drinking, smoking, adjusting radio, adjusting environmental control, reaching for objects in car, and cell phone use. The report stated that 5,474 people were killed and 448,000 individuals were injured in motor vehicle crashes involving distracted drivers in 2009. Approximately 995 deaths of those individuals were drivers distracted by cell phones. In 2013, 21,132 occupants died in motor vehicle traffic crashes. Of the 21,132 passenger vehicle occupants killed, 9,777 were known to be restrained. Restraint use was not known for 1,775 of the occupants. Looking at only occupants where the restraint status was known, 49 percent were unrestrained at the time of the crash.

The NHTSA states that 80% of accidents and 16% of highway deaths are the result of distracted drivers. The National Safety Council (NSC) estimates that 1.6 million (25%) of crashes annually are due to cell phone use, and another 1 million (18%) traffic accidents are due to text messaging while driving. These numbers equate to one accident every 24 seconds attributed to distracted driving by cell phone use. The NSC also reported that speaking on a cell phone while driving reduces focus on the road and the act of driving by 37%, irrespective of hands-free cell phone operation. The US Department of Transportation estimates that reaching for a cell phone distracts a driver for 4.6 seconds, or the equivalent of the length of a football field, if the vehicle is traveling 55 miles per hour. It has been shown that reaching for something inside the vehicle increases the accident risk by 9 times. Texting while driving increases the risk of an auto accident by 23 times.

A 2003 study of U.S. crash data states that driver inattention is estimated to be a factor in 20–50 percent of all police-reported crashes. Driver distraction has been determined to be a contributing factor in estimated 8–13 percent of all vehicle crashes. Of distraction-related accidents, cell phone use may range from 1.5 to 5 percent of contributing factors, according to a 2003 study.

"Outside person, object, or event" (commonly known as rubbernecking) is the most reported cause of distraction related accidents, followed by "adjusting radio/cassette player/CD." "Using/dialing cell phone" is the eighth most reported cause of distraction-related accidents, according to the study.

According to the article "NHTSA distracted driving guidelines" in the August 2013 Motor Age magazine issue, the NHTSA released voluntary guidelines covering the use of in-car infotainment and communication devices, that have some bearing on connected car technologies and In-Car Interactive System. "Proposed items include disabling manual text entry and video-based systems prohibiting the display of text messages, social media or Web pages while the car is in motion or in gear. The goal: Don't take the driver's eyes off the road for more than two seconds at a time, or 12 seconds in total by limiting drivers to six inputs or touches to the screen in 12 seconds". In 2011, according to the NHTSA, 1/3 of the accidents caused by distracted driving. Driving and eating is very distracting. A correspondent for the Boston Globe, Lucia Huntington, stated, "Distracted driving is the cause of many of today's traffic accidents. In a world of ever-extending commutes and busy schedules, eating while operating a vehicle has become the norm, but eating while behind the wheel proves costly for many drivers. Soups, unwieldy burgers, and hot drinks can make steering a car impossible. Although the danger of eating while driving are apparent and well known, drivers ignore them repeatedly, accounting for many crashes and near-misses." During a study done by NHTSA, the NHTSA blames "inattentive driving" for 80% of all car accidents. 2.1 percent of the total were daydreaming, personal hygiene, and eating. Location is another factor to be considered, now that people are living in the suburbs, this has caused a longer commute to work for some.

A study by Monash university found that having one or more children in the car was 12 times more distracting than talking on a mobile phone while driving [22]. According to David Petrie of the Huntington Post, children in the back seat are the worst distraction for drivers. While the focus on texting while driving is laudable, it has failed to address long-standing issues. In both cases an incoming call and a crying child create a situation where the driver should pull over and

not attempt to multitask. A study by AAA found that talking to a passenger was as distracting as talking on a hands-free mobile phone.

More than 600 parents and caregivers were surveyed in two Michigan emergency rooms while their children, ages 1–12 years were being treated. During this survey, almost 90% of parents reported to be engaged in at least one technology-related distraction while driving their children in the past month. The parents who disclosed using the phone—hand held or hands free—while driving were 2.6 times likely to have reportedly been involved in a motor vehicle crash. The rising annual rate of fatalities from distracted driving corresponds to both the number of cell phone subscriptions per capita, as well as the average number of text messages per month. From 2009 to 2011, the amount of text messages sent increased by nearly 50%.

Driver Behavior Detection and Support

A promising strategy to minimize the effect of distraction is to develop intelligent in-vehicle systems, namely adaptive distraction mitigation systems, which can provide real-time assistance or retrospective feedback to mitigate distraction based on driver state/behavior, as well as the traffic. Such systems must accurately and non-intrusively detect whether drivers are distracted or not. Driver distraction detection is nothing but comparison of driver behavior (1) in the normal driving without distraction and (2) driving with distraction. Visual distraction relates to the driver's eyes-off-road time. A general algorithm that considers driver glance behavior across a relatively short period could detect visual distraction consistently across drivers. Detecting cognitive distraction is much more complex than visual distraction because the signs of cognitive distraction are usually not readily apparent, are unlikely to be described by a simple linear relationship, and can vary across drivers. There are four ways to measure driver distraction:

Driver Biological Measures

Cognitive distraction can be measured through a variety of physiological techniques. One such promising approach is to use signals of Electroencephalographic (EEG) activity, referred to as Event-Related Brain Potentials (ERPs). This technique provides a window into the brain activity

that is associated with responses to imperative driving events (e.g., brake lights on a lead vehicle). Using this technique, it was founded that the brain activity associated with processing the information necessary for the safe operation of a motor vehicle was suppressed when drivers were talking on a cell phone. However, this method is frowned upon due to its intrusive nature for real-time scenarios [23].

Driver Physical Measures

The most commonly used driver physical data for driver cognitive distraction are eye movements. Research has proved that cognitive distraction causes drivers to concentrate their gaze in the center of the driving scene, as defined by the horizontal and vertical standard deviation of gaze distribution, and diminishes drivers' ability to detect targets across the entire driving scene. In human science and psychology studies, it has been proved that mouth movement is a good indicator of a human's state of mind and when a person is thinking, his/her mouth and eyes are moving together. Mouth movement can also be thought of a form of body language. Two important conclusions from their study are: mouth and eye movements are highly correlated to each other; and right eye is more correlated to mouth movement either from eye's height or width compared to the left eye.

Driving Performance Measures

A change in the mental state can induce the change in driving performance. Many studies prove the fact that compared to the attentive drivers the distracted ones steer their car in a different way; the same applies for throttle use and speed. Some lines of evidence show that drivers adjust their behavior according to cognitive demand of secondary tasks. Drivers tend to increase the headway distance when they engage in cognitively demanding secondary tasks. This suggests that drivers may compensate for the impairments that secondary tasks have imposed as elaborated earlier. A study introduced a technique for online driver distraction detection that used LSTM (Long Short Term Memory) recurrent neural nets to continuously predict the driver's state based on driving and head-tracking data. The measured signals include steering wheel angle,

throttle position, speed, heading angle, lateral deviation, and head rotation. These links between driving performance and cognitive state show that driving performance measures are good candidates to predict cognitive distraction [24].

Hybrid Measures

In one of the above study, driver physical measures and driving performance measures were combined to detect driver distraction in real time. Machine-learning techniques were used to detect driver cognitive distraction based on the standard deviations of eye gaze, head orientation, pupil diameter, and average heart rate. Sathyanarayana et al. detected distraction by combining motion signals from the leg and head with driving performance signals using a k-nearest neighbor classifier, the driving performance signals adopted including vehicle speed, braking, acceleration, and steering angle.

Among all of these measures, eye movements are one of the most promising ways to assess driver distraction. There are limits in the process of extracting eye movements' parameters [24]:

- Complex calibration: Before each drive, the calibration of the gaze vector with the screen must be verified according to each driver's height and position.
- Driver restriction: The participants cannot wear sun glasses or eye make-up because these conditions can negatively affect tracking accuracy.
- Environmental restriction: Eye trackers may lose tracking accuracy when vehicles are traveling on rough/bumpy roads or during improper lighting conditions.
- Time delay: The software/system will take at the least few seconds to transfer camera image and analyze.

U.S. DOT Recommendations to Minimize In-Vehicle Distractions

The guidelines include recommendations to [10]:

- Limit the time a driver must take his eyes off the road to perform any task to 2 seconds at a time and twelve seconds total.

- Disabling several operations unless the vehicle is stopped and in park, such as:
 - Manual text entry for the purposes of text messaging and internet browsing;
 - Video-based entertainment and communications like video phoning or video conferencing;
 - Display of certain types of text, including text messages, web pages, social media content.

CHAPTER 3
RESEARCH APPROACH

We believe in defining a process before starting with the implementation because the process will allow carrying out work between the different engineering teams of a project. Our research approach starts with a general analysis of the product in terms of engineering, including the deep market research of the hardware and software. Second step would be the domain analysis with a focus on the technology capabilities and limitations. This is followed by an analysis of close and remote competitors. And finally identification of the future trends, looking at the trend where all the manufacturers are focusing.

After the analysis, start with the design, development and testing of the prototype. There can be multiple iterations of prototyping based on the result analysis. Lastly, the aim is to deliver a product which has been tried, tested and true along with proper documentation.

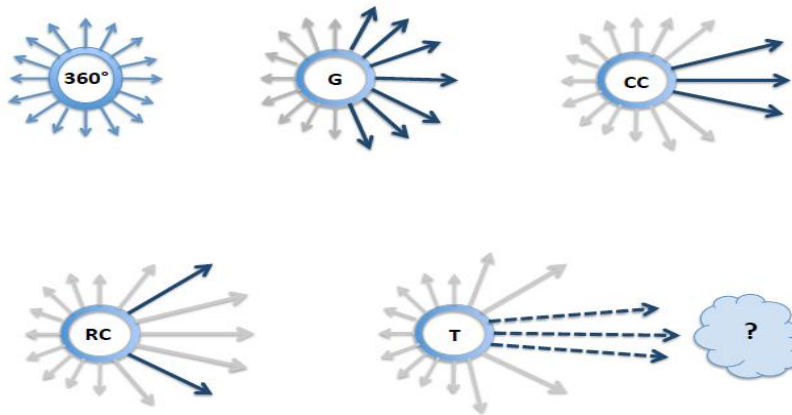


Figure 4: Progressive Domain Analysis

This process has to be followed at every step within every team including the electrical, mechanical or software engineering teams.

Process Implemented in EcoCar3 project

This research approach is being followed in the Innovation Team of an ongoing project EcoCar 3. The work was presented in first week of June 2015 as a part of Year 1 workshop, Seattle.

Project Goal: EcoCAR 3 is the U.S. Department of Energy (DOE) Advanced Vehicle Technology Competition (AVTC) series. EcoCAR 3 is challenging 16 North American university teams to re-design a Chevrolet Camaro to reduce its environmental impact, while maintaining the muscle and performance expected from this iconic American car.

Goal of Innovation Team: To reduce driver distraction and increase driver safety by implementing multimodal enhanced human-car interaction.

Figure 4 is the visual representation of our innovation process. The first step is 360 degree analysis that included the research on automotive domain. Followed by general analysis of in-car voice recognition and ADAS. We have analyzed about 400 images of the car cockpits from the different car models and the different car categories like economy, premium, muscle, sports, and luxury. After having holistic view of automotive domain.



1960 Dodge Charger



2000 Dodge Charger



1960 Chevrolet Camaro



2000 Chevrolet Camaro

Figure 5: Trends in Muscle Car

Third step, the close competitor's (CC) analysis of the three American muscle cars: Chevrolet Camaro, Dodge Challenger and Ford Mustang. Followed by the remote competitor's (RC) analysis of Mercedes SL class and BMW. Lastly, identified the future trends (T). Example as shown in figure 5, the body of muscle cars was sharp angled and in late 90s it started molding into rounded edges.

The same process has also been started in electrical team of ecocar3 with a focus the details under the dashboard.

Today's Car: Computer on Wheels

Today, most of the devices are computerized. The car is no longer just a vehicle powered by an engine, rather it has become more like a computer on wheels. The computer software is needed from the design to the manufacturing to the runtime.



Figure 6: Computer-on-wheels

Design and Modeling Time

Use of software products such as AUTODESK tool specially designed for automotive manufacturers and suppliers help engineers and designers their 3D CAD designs virtually, before they are built; SOLIDWORKS to integrate 3D designs with simulation for optimization, Siemens PLM Software's product NX is a CAD/CAM/CAE software suite etc.

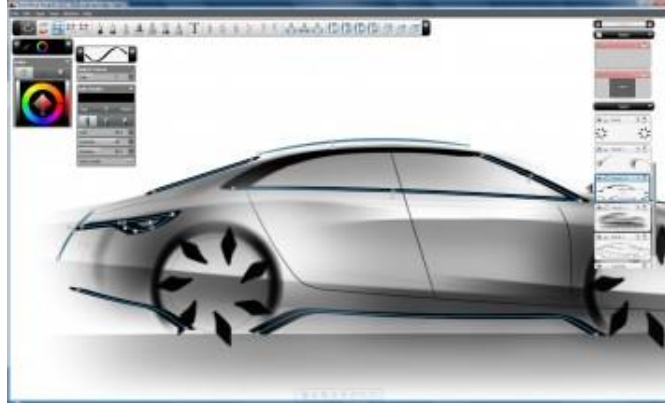


Figure 7: Autodesk Tool for car design

Figure 7 shows the designing of a car using Autodesk tool before the manufacturing. Figure 8 shows the car simulator designed using CarMaker to monitor the working and processing of vehicle before running it on the roads.

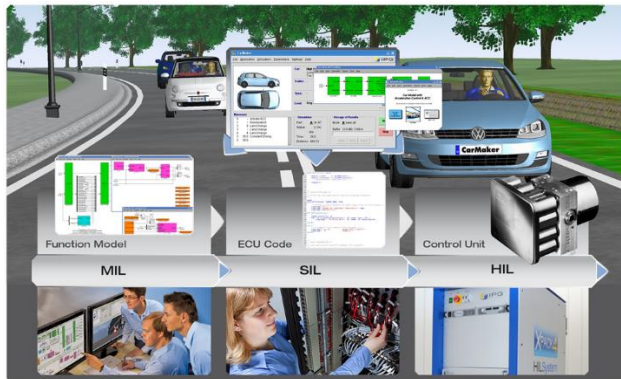


Figure 8: CarMaker for Simulink for Model-Based Design, development and testing of cars

Build Time

To keep up the supply according to the demand, auto manufacturers opted a solution of robotic automation for manufacturing. Automated robotic systems can be used to perform all kinds of automotive tasks on vehicle frames, fenders and underbodies, as well as other parts of the vehicle during production. While changing a line to work on a different car model used to be tedious and slowed production, automated robots today can be reset without having to be reprogrammed each time -- without stopping the line, which increases versatility and productivity for manufactur-

ers. In the late 1990s and early 2000s, Chrysler started robotic automation processes in their factories. Ford also started the robot automation train, dumping \$1 billion dollars into automating their 50+ year old factory over the last decade.



Figure 9: Nissan's Car Manufacturing plant

Runtime

This is our major area of concern where the computer has most influence on the car. Most parts of the car such as the engine, suspension, controls systems and car cockpit controls including infotainment and ADAS are monitored and controlled by computer software. The processors provide a ready source of power, ventilation, and mounting space and sell in terrific quantities. The 7-Series BMW and S-class Mercedes have about 100 processors each. Even a comparatively low-profile Volvo still has 50 to 60 baby processors on board. The first car to use a microprocessor was the 1978 Cadillac Seville [25].



Figure 10: Mercedes S class

Advance Driver Assistance Systems (ADAS)

ADAS in smart cars help the driver to make driving easier and safer. The challenge is to make the driver aware of the assistive technology and its features in the car while providing an easy way to interact with them without getting distracted.

In context to the multimodal interaction, it is required that we virtualize all of the available Manual button related to ADAS as well. The broad classification of ADAS systems presented by DERSEV is given in this section. It is classified in 10 groups with each group having several applications that are currently available or will be soon introduced in the automotive market. The classification of ADAS is as follows:

1. Lane Change Assistance Systems: This category of ADAS includes Lane Departure Warning System (LDWS), Lane Change Assistance System (LCAS), Overtaking Assistance System, and Blind Spot Detection (BSD).

- Lane Departure Warning System (LDWS): signals the driver with acoustical or haptic warnings before the vehicle is about to leave the lane. When the sensors notice that the car is wandering across the lane markings and the indicators are not in use, typically a computer sends a signal to a pair of vibration devices, on each side of the driver's seat. LDWS are available in many cars today. One example is Audi A4 lane assist system- the steering wheel vibrates once only in order to make the driver aware when the vehicle is approaching or crossing a spotted lane marker. The second warning is given only if the vehicle has moved an adequate distance away from the lane marker.
- Blind Spot Detection (BSD): Generally, a quick look at the inside and outside mirrors, possibly even a momentary glance over the left shoulder, we pull out to overtake and then a major fright happens when there is loud hooting from our left. As we fail to see the car approaching quickly from behind in the left-hand lane or in the blind spot next to our own car easily happens, particularly in a heavy traffic on the multi-lane freeways or highways and in urban traffic as well. The Blind Spot Detection System (BSDS) can monitor this area and take much of the worry off the driver and avoid dangerous situations. Blind

spot detection warns the driver about cars that are approaching from the rear or cars that the driver is currently overtaking.

2. Forward or Rearward looking Systems: includes Collision Warning System, Low Speed Collision Avoidance System, Pre Safe System, Collision Avoidance System, Emergency Braking ahead, Electronic Emergency Brake Light, Intelligent Intersection (Emergency Vehicle Detection), Rear Approaching Vehicle, End-Of-Tail Congestion Warning. Collision warning and avoidance is a set of direct supports to the driver to assist safer driving. It covers two distinct sets of applications:

- Collision Warning Systems- Pedestrian Detection System (PSD): provides information about possible collision to the driver, but it remains up to the driver whether to use that information and what action to take. Pedestrian Detection System supports drivers to identify a person near or on the road. These systems have to work in all weather conditions and at night. Also, they must be potent enough to differentiate pedestrians from other objects near the road. One example is BMW Pedestrian Warning system. If a pedestrian come in the car's path, the driver receives an audible and visual important warning in the instrument cluster. Similar, these systems are present in Mercedes S class (Night view assist plus, for the pedestrian and the large animals detection) and in new Volvo series (new Volvo V40, S60, V60, XC60, V70, XC70 and S80) with cyclist detection technology.
- Collision Avoidance Systems: - Emergency brake assistance (EBA): These systems activates an avoidance reaction (e.g. deceleration) when a latent collision is detected. The majority of all rear-end collisions could be circumvented or at least, their harshness could be considerably reduced through timely braking. If the car approaches an obstacle (stationary or moving) and the driver does not react, a warning light activates and is reflected in the windscreen. At the same time, an audible buzzer sounds and a brake function is automatically activated to build up higher braking pressure. The EBA feature is also available in the different configurations. The rear-end collisions mostly occur in inter-urban areas. The EBA-City, an entry-level version, can prevent accidents in these areas

at speeds of up to 25 km/h. The above described function can be realized with the technologies such as Multi-Function Camera with Lidar, Short Range Lidar and Rear Cross Traffic Alert.

3. Adaptive Cruise Control (ACC): can not only maintain the speed chosen by the driver, but also monitors and maintains the headway distance. While driving at a lower speed, the moment another vehicle ahead is within a certain distance, long range radar mounted in the front detects the situation and ACC adjusts the distance by braking the car the exact amount that's needed when activated, ACC give gas and to some extent applies the brake in a way to keep as high comfort as possible.

4. Adaptive Light Control Systems: includes at the moment Adaptive High Beam Assist, Inter Urban Light Assist, Map supported Frontal Lighting, Partial High Beam Assist. A light-beam controller is used to support drivers in controlling vehicle's beams increasing its correct use, since usually drivers do not switch between high beams and low beams or vice versa when required. The adaptive light controller manages the spinning modules so that they always provide the perfect light for interurban, urban and highway driving.

5. Park Assistance System (PAS): helps drivers in parking their vehicle via an in dash screen and button controls. The car can navigate itself into a parking space with slight input from the driver. The first solution in the market had been introduced by Toyota. In the Toyota Lexus system, the driver is accountable for checking if the symbolic box on the screen correctly recognizes the parking space. If the space is large enough to park, the box will be green in color; if the box is incorrectly placed, or lined in red, using the arrow buttons moves the box until it turns green. Once the parking space is correctly identified, the driver confirms and take his/her hands off the steering wheel, while keeping the foot on the brake pedal.

6. Night Vision System (NVS): Anything that generates heat such as a person, an animal and to some extent trees and bushes can easily be monitored on the display. NVS makes it possible for the driver to discover an object much sooner. The system can be also found in cars like BMW,

and Cadillac. Thanks to an infrared camera, mounted in the front of the car, the driver can when driving in the dark, discover a human being or an animal up to 300 meters away.

7. Traffic Sign and Traffic Light Recognition Systems

- Traffic Sign Recognition System (TSRS): As we know, a failure to see a road sign displaying the permissible maximum speed can be expensive in terms of money as well as life. Traffic Sign Recognition System (TSRS) has a display on the instrument panel to remind drivers of the current speed limit. Currently, this system is also available in the Volkswagen Phaeton and in several Volvo models. This is achieved through multiple use of the same camera which is also used for the Lane Departure Warning system.
- Traffic Light Recognition System (TLRS): The system pass on the traffic light information to vehicle, providing alerts to the vehicle occupants via the audio system and on-screen on the navigation system. For instance the BMW traffic sign recognition system depicts overtaking ban or speed limit on the instrument panel in the form of a traffic sign until the restrictions is changed or lifted. In the Mercedes S class 2014 solution a visual and acoustic warning is additionally output in the instrument cluster.

8. Navigation and Map Supported Systems

Includes the Curve Speed Warning System, designed to avoid drivers from entering a curve at a speed faster than the speed permissible at the impending part of the route. On every occasion the driver go beyond this critical speed, a warning is given out.

9. Vehicle Interior Observation and Driver Monitoring Systems

These systems include driver impairment warning system (e.g. drowsiness, fatigue), driver visual distraction warning system (e.g. focus on the driving task, eye gaze evaluation), occupant detection system. When a driver monitoring system detects signs of driver fatigue or drowsiness, the systems will typically start off by sounding some type of buzzer or chime and illuminating a light

on the dash. If the driver stops driving erratically at that point, the system will typically shut off the nag light and reset itself. However, if the signs of fatigued driving continue, the driver alert system may sound a louder alarm that requires some sort of driver interaction to cancel. Some driver alert systems eventually progress to an alarm that can only be cancelled by pulling the vehicle over and either opening the driver's door or shutting the engine off.

10. Autonomous Driving

Some of the autonomous technology is listed below, which are similar to ADAS in context to assisting the driver in different modes.

- **Low Speed Companion:** takes control of braking, starting, and adherence to a safe following distance, leaving the driver free. Through connectivity with the infrastructure, the vehicle even recognizes when the traffic jam comes to an end, reliably turning these functions back over to the driver.
- **Parking Companion:** feature lets any driver easily conquer any parking space. Once the assistant function is activated, the vehicle automatically scans parking areas for a suitable space while passing by and then offers that space to the driver.
- **Parking Pilot:** In this process, the vehicle is operated via a special smartphone app, for example. The driver initiates the parking process after leaving the vehicle. The vehicle connects with the infrastructure – such as the parking lot – and drives to an assigned parking space completely automatically. When the driver wishes to move on, the vehicle is then called back up using the smartphone.
- **Highway Chauffeur:** feature allows the driver to relax and remain inattentive as the vehicle handles all of the management-related tasks, securely overtaking slower vehicles and even conquering complex situations, such as changing highways, driving in tunnels, and toll booths.
- **Highway Pilot:** provides all the features of highway chauffeur as well as an additional safety feature that allows driver to remain inattentive to the traffic conditions. In any

emergency situation that may occur, the vehicle will be able to automatically pull over and place an emergency call to ask for help.

CHAPTER 4

AUTONOMOUS VEHICLES

The rapid advances in information communication and technology are now incorporated into the transportation system to overcome challenges faced in terms of management efficiency and safety of various modes of transport. Intelligent transportation system (ITS) is multidisciplinary and its research activities are spread over different areas electronics, sensors, information systems, robotics, and communication systems.

The five major components of intelligent transport system are [26], [27]:

- **Advanced Traffic Management Systems (ATMS):** use advanced technology to monitor traffic conditions and provide real-time solutions. The main elements of ATMS are data collection, support systems such as sensors, cameras, and electronic displays etc., real time traffic control systems that use the data collected by other two elements.
- **Advanced Traveler Information Systems (ATIS):** The travel from one location to another is made easy if the drivers able to plan their trip based on the constant traffic updates and best routes possible to choose destination. The information collected from ATMS is feed on to ATIS will help in reducing the amount of delay in travel and congestion problems.
- **Advanced Vehicle Control Systems (AVCS):** With present day computing power, control systems and sensor technology helps in creating safer mobility to people. Sensor technology is used as the visual aid to the drivers and information source other vehicular systems. Automated controls such as braking, acceleration, advanced cruise control, automated steering handy for drivers.
- **Commercial Vehicle Operations (CVO):** It is majorly aimed at using advanced technologies to increase the safety and efficiency of commercial vehicles and fleets.

- Advanced Public Transportation Systems (APTS): APTS uses ATMS, ATIS and AVCS technologies to improve the mass transportation systems. Improving the accuracy in terms real time traffic information, routes, means of available transport and schedules.

Of all the above areas Advanced Vehicle Control Systems would be one of the most researched area in all possible modes of transport. The state of wireless communication technologies, sensors, lightweight materials has resulted in considerable automation in transportation systems. The automation is being expanded further to create autonomous vehicles. Autopilot in aircraft cruising, Drone - common word for Unmanned Aerial Vehicle, Auto driving/Driverless trains used in urban subways and Autonomous road vehicle – Google Driverless Car are the few major developments in this direction.

Autopilot

Autopilot as the name suggests that the airplane flies without the need of human pilots. It can almost fly the plane completely between takeoff and landing. The autopilot system relies on a series of sensors around the aircraft that pick up information like speed, altitude and turbulence. That data are ingested into the computer, which then makes the necessary changes. Basically, it can do almost everything a pilot can do. Apparently, it is "The autopilot system does not fly the airplane, but the pilots fly the plane through automation".

In recent years, pilot interaction with automated flight control systems has become a major concern in the transport industry. This problem has variously been termed as lack of mode-awareness, mode-confusion, or automation-surprises. The three most significant phases of flight are Takeoff, cruising and landing. Cruising the majority part of journey where the altitude, speed remain constant, most of the time change is only observed with respect to the direction of flight. Autopilot was introduced to relieve the pilots of this mundane job and reducing the workload on pilot in terms of guiding and controlling the flight. It is used independently to maintain the speed and altitude of the flight or along with the navigation system [28].

The interface between the user and the machine provides inadequate information about the status of the machine [29]–[31]. The user has an inadequate “mental model” of the machine’s behavior [32]–[34]. In high risk systems, such as commercial aviation, faulty interaction of the user with the machine has led to catastrophic results [35]. This faulty interaction has been variously attributed to a combination of human and machine problems. However, the distinction between human error, inadequate training, lack of situation awareness, and interface design errors is blurred. One aspect is the complexity of automatic control systems and the lack of rigorous methods for their systematic analysis and evaluation [36].

There are several modes that defines specific behavior. The pilot must interact with this autopilot to perform certain task. For example if a flight has to climb upto 15,000 feet; maintain an altitude of 15,000 feet and later descend to 12,000 feet.

Pilots are also less likely to recognize growing problems in the airplane’s equipment if not periodically engaged. Active monitoring may be difficult to achieve in aviation, given the degree of automation already present in cockpits. But industries in which automation is nascent—automotive, medical, housing construction—still have the opportunity to learn from the problems that have occurred in the cockpit.

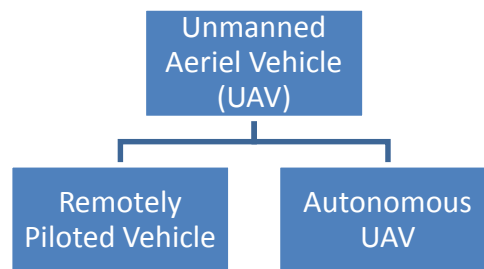
Present-day automated control systems are very complex and user-interaction is, inevitably, complex as well; there is no escape from this reality, at least by today's standards.

Even if we shape up fluent autopilots with the necessary situational awareness we don’t want to perform on autopilot all the time. The Federal Aviation Administration (FAA) put out a recommendation that pilots spend more time in manual control of their aircraft. Their recommendation is based on the concern that pilots’ skills may be degraded because autopilot does not reinforce the pilot skills necessary for manual flight operations, especially if the airplane is in an upset state.

Drones

Drone is a general term used for early Unmanned Aerial Vehicle (UAV), a flying machine that is remotely operated either autonomously by onboard computers or by the remote control of a pilot on the ground or in another vehicle [37]. The UAVs are increasingly autonomous which follow a preprogrammed machine/target reducing the intensity of damage that may occur in case of manned flight. Drones fall into the nomenclature used for categorizing aerial vehicles in military. Drones are majorly used for reconnaissance and surveillance purposes, carrying ammunitions for military striking purposes. Although drones are mainly used for military applications there has been unprecedented growth with respect to various civil application domains. Surveillance and reconnaissance are the major markets for UAVs.

The below image shows different type of UAV's



Remotely Piloted Vehicle (RPV):

The pioneers Predators, Scout and Jindivik are the major examples for the Multi Mission Remotely Piloted Vehicle. These are mainly used for armed reconnaissance, surveillance and precision targeting. The equipment and operations are taken care of by two crew member located at the ground control station (GCS). One for piloting the aircraft and other for operation sensors and weapons as and when required. Target drones, decoys and Missiles are remotely piloted aerial vehicles designed for specific purposes. All these aerial vehicles are operated remotely from the launching vehicles [38].

Autonomous UAV Multi Mission:

Global Hawk flies autonomously from takeoff to landing. It proved its capabilities which include high altitude long duration flights along with the capability to produce high quality images [38]. There is a ground station for controlling and monitoring, making operational changes if any. MCE is used for mission planning, control and command, image processing and dissemination. LRE is used for controlling launch of recovery and associate ground support equipment. Ground segments are equipped with antennas for Line Of Sight (LOS) and satellite communications with air vehicles [19]. Cruise missiles and Decoys are the examples of Autonomous UAV which are directed at a specific task.

It can be seen that there is major human involvement in case of remotely piloted aerial vehicles. Although there is no specific classification human involvement it varies from one UAV to another depending on the level of automation. This leads to study of human factors that have caused the UAV mishaps. The current UAV mishap rate is 1 for 100,000 flight hours and this has to be significantly reduced to meet the safety standards. The issues related to UAVs usage in Army, Navy and Air force differed significantly. In sum, Air force UAV failures involved instrumentation and sensory feedback systems, automation and channelized attention, Army mishaps can be attributed to procedural guidance and publications, organizational training issues and programs, operator overconfidence, and crew coordination and communication. Navy/ Marine failures were due to workload, attention and risk management. These can be broadly categorized as organizational influences, preconditions, supervision and act based mishaps.

Applications of UAVs include but not confined to cartography, border patrol, pipeline patrol, port security, drug surveillance, traffic monitoring, inspection, homeland security, search and rescue, fire detection, agriculture imaging, land use mapping, flood mapping and imaging. There are small-scale unmanned aerial vehicles called "Microdrones" used for providing bird's eye view of the environment. Unmanned aerial vehicles are advantageous over manned vehicles in terms of safety, cost reduction, flexibility in planning and reaching the target. The data acquired in terms of images has high resolution and can be used to provide comprehensive information to other controlling systems.

Even with such high prospects there are issues deterring integration of UAVs in civil and commercial domains. Notable among them are immaturity/ lack of consensus in concepts related to classification, concepts and definitions, certification standards for UAV systems, operator training, Limited payload capacity, space restrictions, safety standards, cost liability for acquiring the systems and difficulties related to frequency spectrum availability. Currently armed drones are used by US, United Kingdom, Israel, China and Russia [38].

In conclusion it can be said that drones have huge potential in future if the issues pertaining to commercial and civil integration along with safety and privacy concerns can be addressed. UAVs act as the part of complex system. The future would be the development of fault tolerance, collision avoidance and reconfigurable systems which can perform complex tasks at ease. Main challenge would be dealing with networked UAVs with reliable autonomy.

Subways

This concept is related to the automated metro lines, which are now implemented in most of the cities. The benefits of automated metro lines are that they are cost effective and help in high traffic movement. There are three different levels of automation possible in metros [35]:

- Semiautomatic train: There will always be a driver in the cockpit. Driver is responsible for departure from the current station and the next station is reached without his intervention. Driver is solely responsible for initiating the necessary steps during emergencies.
- Unattended train: All the train operations are overseen by array of remote technologies, from CCTV and onboard telemetry to automatic detection systems. These systems are responsible for detection and management of all possible hazardous situations. There are no on-board staffs.
- Driverless train: This technology requires automatic handling of all aspects of the operation which means that there is no driver sitting in the cockpit. Dubai Metro, Tokyo, Paris and Barcelona are embracing true driverless systems, with trains running automatically.

There may or may not be any operation staff present in the train depending on the system maturity.

In a fully automated metro the first major change is that there is no driver in the train's cockpit. Instead there are on-board computers for controlling the train and human operators are situated in remote control rooms supervising the traffic situations. The tasks of driver in a train includes driving preparations, answering emergency calls on train from control station, driving the train on the track, stop the train at the station, reacting to the events that occur around, opening and closing the train doors, changing the direction of trains at terminals, reacting to exceptional situations, keeping the train on schedule, contacting Traffic control, making announcements to the passengers etc. The automated control system should be capable of performing all these tasks similar to the driver. These Driverless trains/ automated control systems are reported to be success in the Bay Area Regional Transit system in San Francisco, Orlando, Tampa Bay and other places.

The major concern with driverless trains is safety. But the operation environment for trains is certainly different from other means of transport. Given the fact that path and distance between the stations is always fixed trains can be easily automated compared to other means of transport. The Advanced communication based control systems are pointing towards the future of train transportation infrastructure which is capable of unattended train operation.

Automated Road Vehicle- Google Car

Over the past few years the automobile industries are making advancements in bringing digitalization and automation into exclusively a human function: driving. There are many new car models increasingly rapidly with semi-automatic features such as adaptive cruise control, on board navigation and collision avoidance, and parking assist systems that allow cars to steer themselves into parking spaces. Few organizations started inventing autonomous vehicles that can drive themselves on existing roads and navigate on roads and environmental contexts with almost no human presence. Many automobile industries, like Audi, BMW, Cadillac, Ford, GM, Mercedes-Benz, Nissan, Toyota, Volkswagen, and Volvo are researching and running test runs on automated vehicles.

NHTSA defines vehicle automation as having five levels [39] :

No-Automation (Level 0): The driver completely controls the vehicle – brake, steering, throttle, and motive power – at all times.

Function-specific Automation (Level 1): involves one or more specific control functions such as electronic stability control or pre-charged brakes, where the vehicle automatically assists with braking to enable the driver to regain control of the vehicle or stop faster than possible by acting alone.

Combined Function Automation (Level 2): involves automation of at least to primary control functions designed to work in unison to relieve the driver of control of those functions. An example of combined functions enabling a Level 2 system is adaptive cruise control in combination with lane centering.

Limited Self-Driving Automation (Level 3): The driver cedes full control to the vehicle. The driver is expected to be available for occasional control, but with sufficiently comfortable transition time.

Full Self-Driving Automation (Level 4): The vehicle is designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip. Such a design anticipates that the driver will provide destination or navigation input, but is not expected to be available for control at any time during the trip. This includes both occupied and unoccupied vehicles.

Automated Vehicles have the potential to modify transportation systems by turn away crashes, providing critical mobility to the elderly and disabled, increasing road capacity, saving fuel, and lowering emissions. Complementary trends in shared rides and vehicles may lead us from vehicles as an owned product to an on-demand service. Infrastructure investments and operational improvements, travel choices and parking needs, land use patterns, and trucking and other activities may be affected. One such automated vehicle is Google Car developed by Google.

Google Car consists of GPS, stored maps, LIDAR (Light Detection and Ranging), and a set of onboard cameras in the vehicle [40]. The first two of these technologies allow the vehicle to un-

derstand where it is in the world where it should be going, and how to get there; the latter two allow the vehicle to track where it is on the road and where other vehicles, traffic indicators, and pedestrians are. While termed “driverless,” the vehicles are better classified as driver-optional, where human operators are expected to have either primary or secondary control responsibilities.

Benefits [41]:

The critical reason behind a crash is comprises of the vehicle, environment, additional human factors such as inattention, distraction, or speeding are regularly found to have contributed to the crash occurrence and/or injury severity. These vehicles have capability of avoiding crashes, as they are programmed to not break traffic rules, they are not drink and drive. They reaction times are quicker and optimized to normal traffic. While many driving situations are relatively easy for an autonomous vehicle to handle, designing a system that can perform safely in nearly every situation is challenging

Apart from making automobiles safer, researchers are also developing ways for technology to reduce congestion and fuel consumption. Technology that allows for smoother braking , fine speed adjustments of following vehicles, leads to fuel consumption reduction, less brake wear. These vehicles are also expected to use existing lanes and intersections more efficiently through shorter headways and more efficient route choices. Many of these features, such as adaptive cruise control (ACC), are already being integrated into automobiles and some of the benefits will be realized before AVs are fully operational.

Safety and congestion-reducing impacts have potential to create significant changes in travel behavior. For example, AVs may provide mobility for those too young to drive, the elderly and the disabled, thus generating new roadway capacity demands. Parking patterns could change. It is possible that already-congested traffic patterns and other roadway infrastructure will be negatively affected, due to increased trip-making. Freight transport on and off the road will also be impacted.

Plenty of road tests have been done on these vehicles in San Francisco. However, there were 12 minor traffic accidents with Google car but Google argued that it was driven manually at that point of time [40]. As stated by Erico Guizzo , before sending the self-driving car on a road test, Google engineers drive along the route one or more times to gather data about the environment. When it's the autonomous vehicle's turn to drive itself, it compares the data it is acquiring to the previously recorded data, an approach that is useful to differentiate pedestrians from stationary objects like poles and mailboxes [42]. This seems to be an unreliable approach for a real-time scenario.

Limitations & Impacts

Self-driving vehicles provide good benefits, but there are few parameters to be considered for implementation such as[43]

- Unemployment [44]: With the evolution of automated vehicles, it leads to loss of work for people who are living only on driving, and results in increased unemployment in this sector.
- Technology Constraints [44]: As we came to know that Google's autonomous car relies on four major technologies in its autonomous operations. Each of these systems is vulnerable in some form or fashion and it is not clear whether any redundancy exists, the car will not be able to function correctly.
 - If the GPS fails, the destination becomes unclear.
 - If the LIDAR fails, detect other nearby cars, pedestrians, etc. fails.
 - If the cameras fails, unable recognize a stop sign or the current color of the traffic light. In addition, it is not clear how much advanced mapping is required by driverless cars and the frequency of map updates that are required to maintain an effective 3D world model by which the onboard computer makes decisions.

Google's own researchers admit that inclement weather and construction areas are something they have yet to master [41]. Precipitation, fog, and dust are known problems for LIDAR sensors,

which can interfere with the image detection capabilities of the camera and can scatter or block the laser beams sent out by the LIDAR. Cameras are also sensitive to such problems. This leaves the vehicle unable to sense the distance to other cars and to recognize stop signs, traffic lights, and pedestrians.

- **Vehicle Costs:** This technology needs new sensors, communication equipment, and software to guide for each automobile. Current civilian autonomous vehicles are costing around \$100,000. Over time with technology advancement and large scale production provides greater affordability. If prices are near to conventional vehicle price, the people may afford to purchase.
- **Licensing:** Each state has different DMV regulatory licensing and provisional testing standards. Without a consistent and set of standard set of safety for acceptance it is difficult for manufactures. Licensed Drivers of one state are legally able to operate vehicle in other states through set of agreements. Current existing laws may probably not permit automated vehicles in state without automated vehicles license. Failure to clarify the regulation discourages their introduction.
- **Liabilities:** A driverless vehicle on road opens up way for insurance and liability issues. Even with perfect automation it cannot be claimed to have crash-free roads. Even though, the sensors, interpretation software and algorithms are the available and provide informed decisions, there is initial perception that they are potentially unsafe with lack of human driver. If vehicles are at high standard than human, the vehicle cost increases. Some steps have to be implemented for liability concerns, method of storing data has to be established before collision so that based on data we can figure out the fault party during claim.
- **Security:** Electronic security is one of the major concerns for all the people. Hackers, dissatisfied employees, anti-social element organizations may target automated vehicles and intelligence systems resulting in traffic concerns and accidents. There may be a situ-

ation where virus can be programmed and get the system affected and raise issues. The Communication system or sensors may be altered and can be made out of control. GPS Spoofing can also be done which leads to false destinations and lead to kidnapping, moving vehicles into buildings, off bridges etc.

- Privacy: Privacy of the users also needs to be taken into account. Privacy is one of the basic rights and not a privilege. Thus, it is the primary responsibility to ensure that autonomous cars do not violate privacy of an individual. Privacy and surveillance issues may arise from the use of GPS and their communication devices. Though the fact that it is not controlled from a central server might indicate that it is not being tracked but considering the fact that while one car drives without driver, it needs to be connected through cellular network to navigate. This will allow it to pick/drop-off passengers at the right place which is the concept of autonomous public transport. The whole system is a big source of private data and at any point it is trivial to pull up the exact location and efficiently track anyone giving rise to a huge privacy concern. There are two standard protocols available for communicating between autonomous vehicles, namely V2V and V2D. In vehicle-to-vehicle (V2V) protocol the vehicle receives and shares its internal data with other similarly connected, data-sharing vehicles². Whereas the Vehicle to Device (V2D) or Vehicle to Infrastructure (V2I) where the vehicle shares its internal data ((speed, velocity, heading, etc.) with Law Enforcement agencies, traffic management centers etc.

The idea of autonomous technology is improving quickly and some features are already offered on current vehicle models as a part of ADAS system. These vehicles will have real and quantifiable benefits but with many limitations having a human driver is preferable to control the worst case scenarios. Although such vehicles are supposedly capable of driving through any traffic situation without requiring a human driver to apply pressure to the pedals, shifting, or steering, this driver may still choose to do so and may play a role in avoiding accidents. Furthermore, Google car can be considered as a human direction of ADAS, automated of human operator as well. However, Google admits that, while testing, there remains a safety driver and a software

operator in the vehicle at all times to take control in the case of near accidents or software failures [45]. Also, interaction between human and driverless car is still unclear [45]. There are various scenarios where human interaction will be required such as providing the destination to the driverless car, recovering from automation errors, taking over the control in case of emergency etc. The multimodal interaction interface approach can also be applied here.

On the basis of current literature review, it can be concluded that all modes of transport are moving towards the autonomous control systems. By automation, we mean the absence of human intervention and it is the highest possible level of automation in vehicles. Autopilot may sometimes be equated to autonomous control, but we can see that there is always pilot who takes control back in case of complex tasks of flight take-off and landing. However, in drones there is no human onboard controlling the flight but there can be a control station on ground or preprogrammed route. The flying process could be considered more of a deterministic environment. Subways or metro's operating with driverless trains also have structured environment, but still require a control station remotely located. Communication based unattended train operation may be a future possibility. In terms road transport, there is a lot of research in the development of autonomous vehicles, especially cars. The other three environments are more deterministic than the road environment. The environment is dynamic, i.e., continually changing, there cannot be a predetermined route or track or traffic for the vehicle. Apparently, a lot of sensors will be required to keep track of traffic signals and the other entities on road. However, despite doing all that we cannot claim that the risk is fully mitigated. Computer assisted surgery is there but robots performing surgery is still not an option.

CHAPTER 5

MOTIVATION FOR IMMI

Speech Recognition

Speech recognition is a process of converting spoken words into text. It is also known as Automatic Speech Recognition (ASR). It has been widely in use since 1970s in various domains such as automotive industry, health care, military, IT support centers, telephony etc.

In 1773, a Russian scientist Christian Kratzenstein produced vowel sounds using resonance tubes connected to organ pipes [46]. In 1930s, Homer Dudley developed a speech synthesizer called the VODER (Voice Operating Demonstrator) [46]. In 1952, Davis, Biddulph and Balashek of Bell Laboratories built a system for digital speech recognition limited to single-speaker and vocabulary of 10 words. In 1960s, Tom Martin and Vintsyuk worked on the concept of adopting a non-uniform time scale used for aligning speech patterns [46]. In late 1960s, Raj Reddy from Stanford University introduced the continuous speech recognition [47]. The technology gained acceptance and shape in early 1970s due the research funded by Advance Research Project Agency in U.S Department of Defense [46], [47]. Hidden Markov Models, n-gram models and neural networks are widely used in speech recognition systems. Since 2010, deep neural networks is used for learning algorithms to train the system.

The number of applications is growing with each day. However despite of many achievements in recognition of spoken words, it's still not perfect for use in all kind of applications. As stated by Moore's law, the number of computational devices will double every year. This development in computational power has led reduction in running time for large algorithms. There are many common research tools such as Carnegie Mellon University Language Model tool kit, Hidden Markov Model Toolkit, Sphinx and Stanford Research Institute Language Modeling. Derived from communications and information theory, stack decoding was subsequently applied to speech-recognition systems [48]. Viterbi search, broadly applied to search alternative hypotheses, derives from dynamic programming in the 1950s and was subsequently used in speech ap-

plications from the 1960s to the 1980s and beyond, from Russia and Japan to the United States and Europe [48].

Hype Cycle

Introduced in 1995, the Gartner Hype Cycle Model is an idealized model of a typical progress and evolution pattern commonly observed for maturity and adoption of technology.

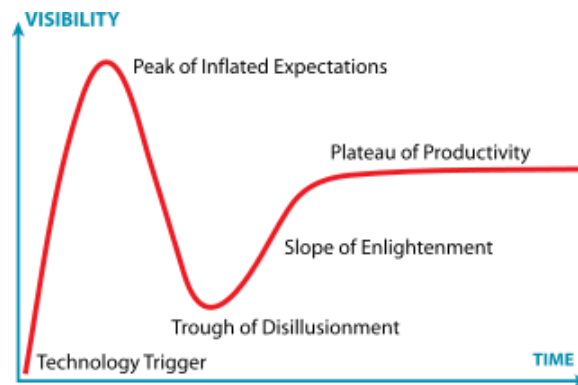


Figure 11: Gartner's Hype Cycle [49]

The Hype Cycle (Figure 11) describes five phases of a technology:

- **Technology Trigger:** is the phase of ideation and conceptualization. There is no actual development but there can be prototyping for proof-of-concept.
- **Peak of Inflated Expectations:** involves publicity about success or failure in first generation production.
- **Trough of Disillusionment:** is the disappointment due to the failure in meeting the expectations of user. Investors continue with the product only if the producers tackle a problem and propose a bug fix.
- **Slope of Enlightenment:** A slower recovery, on the back of second-generation and subsequent products and more companies adopt it.
- **Plateau of Productivity:** where technology becomes more widely adopted. At least 30 percent of the technology's target audience has already adopted it, or is planning to.

The Gartner Hype Cycle for Human-Computer Interaction, 2014, Figure 12, is a good indicator of the systems being developed to let machines understand the user intent rather than just

words[50]. The hype cycle indicates that the SR technology is reaching the Plateau of Productivity in the next year or two.

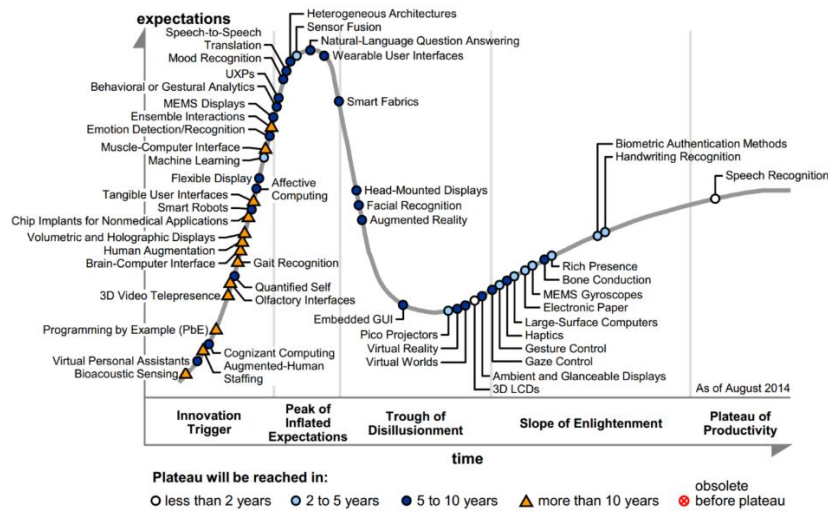


Figure 12: Gartner’s Hype Cycle for Human-Computer Interaction[50]

The Hype cycle is a clear indication of how SR has become an established technology, ready for the market. That could explain why most of the cars come with speech recognition system either as a standard option, or at least as an extra feature. Studies have shown that the total share of new cars with speech recognition is expected to surpass the number of cars without speech recognition by 2020. A survey by Statista shows the twenty percent growth in cars with speech recognition system installed in 2012 and 2019 [51].

Speech Recognition in Automotive Industry

In the USA, current voice interfaces include Ford SYNC, Chrysler UConnect, GM MyLink, Hyundai Genesis, and Toyota navigation with Entune. The commonly supported applications are navigation (e.g., destination entry, route guidance, and traffic information) and music selection (selecting, playing, and pausing songs on MP3players, AM/ FM/XM radios), as well as those related to cellular phones (answering and placing calls, searching contact lists, and various tasks associated with text messages).

Is technology a curse or a blessing?

The use of handsets while driving is illegal in 14 states, whereas the use of “hands-free” voice controls is generally encouraged. Most of the people hold a perception that speech recognition technology is safer because the driver doesn’t need to take eyes off the road or hands off the steering wheel. In contrast, the study by AAA foundation in April 2013 showed that the mental workload from performing complicated tasks slows reaction times, whatever the driver is doing with his or her hands [52].

Speech recognition is assumed to be a panacea for driver distraction offering an alternative to the visual/manual demand of the system. However, speech-based systems can demand attention resulting in increased cognitive load and distracting driver just as visual displays and manual controls do. Past research neither fully supports nor contradicts the above assumption. For example, a research has proved that driver’s emergency braking timings and cognitive workload levels for speech-operative personal assistant devices (PDA) in vehicle is between no PDA and manual PDA [53]. Lee showed that drivers’ reaction time increased by 180ms when using a complex speech-controlled email system (three level soft menus with four-to-seven options for each menu) in comparison with a simpler alternative (three levels of menus with two options per menu) [52]. Another research analyzed the difference in reaction time when no email system is provided in car with the reaction time when provided with speech based in-vehicle email system. The results showed 30% increase in reaction times when used speech-based system [10].

In cars, the environment is brutal. There are a number of disturbances like engine noise, traffic noise, passengers talking, and media player. And hence, it’s still hard to incorporate speech recognition in cars. Anyone who has ever used voice recognition in a car knows that sometimes it is dreadful to use. The most recent annual Initial Quality Study, which focuses on problems new car buyers experience in the first 90 days of ownership, found that 23 percent of reported issues were related to infotainment, and a third of these problems were caused by voice recognition. A separate study by the market research firm J.D. Power & Associates found that the rate of the complaints about built-in voice recognition systems is nearly four times the rate of reported prob-

blems with transmissions or cup holders. In Figure 13, first part show the data about the percentage of a factory-installed voice recognition equipment by brand for 2014 models and second part is the data from J.D Power and associates which shows the number of problems reported per 100 cars. From the research it is clear that so called less distractive hands-free solutions like voice recognition (8.3) and Bluetooth (5.7) are the top most in the complaints received.

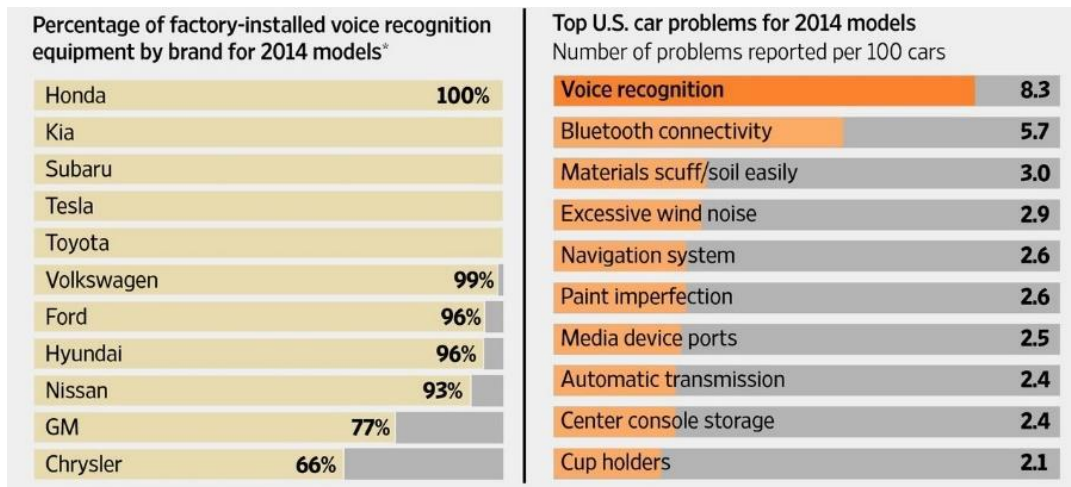


Figure 13: Car Models Made in the North America for U.S Market, 2014

While it would be wrong to say that speech recognition never works. It is due to the engine noise and the noise on road that often makes it inefficient. However, the user interface of center stack also plays an important role when it comes to distraction. The center stack with plethora of buttons can be more distraction than an inefficient speech recognition system. Therefore, working towards achieving perfection for a unimodal system can be difficult, especially when for real-time systems.

A study has proved that the total number of hard and soft buttons at a time on the center stack ranges from 25-50 for the analyzed cars [54]

CHAPTER 6

MULTIMODAL INTERACTION

In simple words we can say, multimodal interaction means more than one means of communication with a system/object to share information. More information can be exchanged in a limited time, often by different means of communication.

- **Sequential Multimodal:** means user will have to switch between the modes of interaction but cannot use the modes together. Example: BT Exact's WAP and voice-based stock quotation application in which the user start with tradition WAP and then switches over to voice to specify the stock in which they are interested [55].
- **Simultaneous Multimodal:** allows user to use more than one mode at a time to interact with the system. For example: in route finder application the user could say "Show me the quickest route from here and here" while indicating the two locations on an on-screen map using touch. No action can be taken based on single mode input but if we combine the inputs, the request can be completed. However, this is not that simple to achieve. The two inputs can be contradictory and may end up in a conflict. There can be various strategies to deal with such situations such as accept only the first input or just the last input, accept most reliable input etc. [55]

Human-Human Multimodal Interaction

If we closely look at the natural way of Humans' Communication, we rarely see dependence on speech as the only source of sharing information. There is an intertwined cooperation of various other modalities [56]. The modality includes a variety of communication methods for the expression of intent, the implementation of action and perception to the feedback, such as speech, eye contact, facial expressions, lip's movements, hands' movements, gesticulation, head's movements, body's posture, touch etc. Even a poor conversation over the phone has multiple modes of communication that work together to help the receiving side easily recognize the speaker.

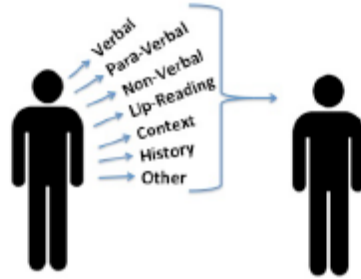


Figure 14: Human-Human Multimodal Communication

Besides the normal verbal communication mode (the actual words used in the conversation), para-verbal modes (like the tone of voice, the pauses, the speed, and the volume of the speaker) are used to enhance recognition (Figure 14). Additional subtle modes like the history related to the conversation as well as the context where it is taking place can also add to better recognition compared to the context free single mode speech. Even for simple text messages or emails, we often use emotion icons as an additional mode to help improve the communication. An advanced face-to-face communication can provide very rich speech recognition as we use additional modes like hand gestures, lip reading and other body- and face impressions.

Human-Computer Interaction

The interface between humans and computers has progressed over the years from switches and LEDs to punched cards, interactive command-line interfaces, and the direct manipulation model of graphical user interfaces. There has been a significant research in multimodal human computer interaction (MMHCI) due to the advancement in the development of unimodal techniques such as speech recognition, computer vision, etc., and in hardware technologies such as inexpensive cameras and sensors. In terms of computer input devices the modalities can be compared to human senses: cameras (sight), haptic sensors (touch), microphones (hearing), olfactory (smell) [57]. In addition, however, there are input devices that do not map directly to human senses: keyboard, mouse, writing tablet, motion input (e.g., the device itself is moved for interaction), and many others. A multimodal framework requires a user-centered approach (rather than a technology-driven approach) to design, engineering, and testing interfaces [58]. One of the early MMHCI

examples is the QuickSet system, an architecture for multimodal integration used for integrating speech and (pen) gestures, allowed users to create and control military simulations from a tablet or handheld computer [58].



Figure 15: Human-Computer Interaction

However, despite the advancement in computer research field, there is still a need for more research on individual modalities such as speech, vision etc. There are challenges in building systems to perform tasks in real-world scenarios such as face detection and recognition, facial expression analysis, hand tracking and modeling, head and body tracking and pose extraction, gesture recognition, and activity analysis. Changes in illumination and camera pose, changes in the appearance of users, and multiple users is a huge challenge for the field. Another important aspect that has been research is emotion tracking. Emotions can change the meaning of message.

CHAPTER 7

IN-VEHICLE MULTIMODAL INTERACTION- OUR APPROACH

Mobile environments rule out many traditional approaches to human-computer interaction (HCI). In order to accommodate a wider range of scenarios, tasks, users, and preferences, interfaces must become more natural, intuitive, adaptive, and unobtrusive.

Multimodal is relatively a new term when used in context of automobiles but has been widely used in Human Computer Interaction. The modes identified for human-human or human-computer interaction are unsuitable for automotive domain due to the dynamically changing environment, the noise and cognitive load on driver. In noisy environments, even humans do not just rely on speech based communication. They tend to gain ideas by facial expressions, body gesture, lip movements etc. Hence, we have identified and experimented with new modes of communication specific to the automotive domain.

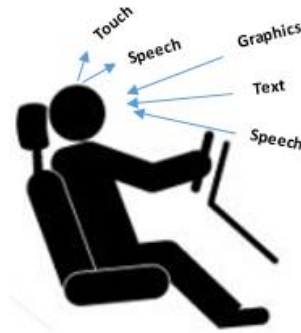


Figure 16: Human-vehicle interaction

For the system-to-driver interaction, we focus on three ways of communications:

- **Graphics:** a visual output mode to represent the exact virtualization of hardware options available. The number of icons is within 6-8 on each screen and graphics are easy to interpret.
- **Text Tip:** A visual output mode to describe the option. Similar to the common “mouse over tip” or “tooltip” text in the mouse-based office interaction, the text tip is a one-word caption to give the user a cue about the voice command that can be used to speak out the command. These one-word captions are used as keywords in speech recognition and

hence, considerably reduces the vocabulary/grammar used to implement system and thereby, enhancing the likelihood of successful command recognition. This mode can be turned on/off by the user based on their preference.

- **Voice-tip:** Each identified command is repeated back to the driver to serve multiple purposes:
 - To provide the user with an idealized way to pronounce the command to improve their future voice-based communications and eliminate the “marginal commands”. Our multiple prototype experiments showed that some users can have the same command recognized some of times but not always. This can be attributed to the distant proximity of pattern matching due to speaker accent or other factors. Hence, if the command was close to the cut-off borders of not being correctly recognized, small perturbations can lead to situations where the command is recognized only some of the time. Therefore, our prototyping with the voice output command shows the user how exemplar command should be spoken, hence improving future user communication with the system.
 - To provide the user with a subtle confirmation, system will speak out the command. By hearing the command that has just been identified and is about to be executed, the user will know if it was the intended command. If not, the user can speak out a “No” or “Cancel” to “cancel” or “undo” the command and repeat or use another mode to reissue it.
 - To provide the user with a voice tip when “Text Tip” mode is off. By touching a particular icon, the system will speak out the voice command that can be used for access. Hence, user will get to know the voice command required for future interactions.

Similar to the Text option, voice output can be turned on/off based on user preference.

For the driver-to-system interaction, we focus on three ways of communications:

- **Touch:** If the user prefers to use the touch screen as their initial interaction mode, s/he can simply choose from a limited number of large icons (6-8 icons/per screen max) to select a feature.
- **Speech:** Daily experience suggests that not all words in the human-human conversation, but only a few important ones, need to be accurately recognized for satisfactory speech communication. Unlike the wide spectrum SR system that listens to a wide variety of commands, the user will have some assistance by the text tip offered on the screen as well as the exemplar commands spoken out in earlier times.

Machine Learning: The issued command is recorded while being identified. If identified successfully, the action will be performed. On the other hand, if the command was miss-diagnosed and the user reverted immediately to the touch mode for another attempt, the original miss-diagnosed command will be paired with the given graphic command and is used in later speech matching. This combined pattern of failed speech followed by a touch command is intended to speed up the system learning process. Also, allows user to change the “keyword/text tip/command” corresponding to any icon as per the preference.

Multimodal interfaces are a natural and safe means of communication and can help in preventing and recovering from the errors due to inaccurate unimodal output. Multimodal interfaces are recognized to be inherently flexible, and can be designed to support simultaneous use of input modes, to permit switching among modes to take advantage of the modality best suited for a task, environment, or user capabilities, and to pass information from one mode to another in order to expand accessibility for users with selective limitations.

Related Research

There is a similar research done on multimodal driver-vehicle interaction, in which speech recognition is combined with facial expression recognition to increase intention recognition accuracy [59]. This would be the ultimate solution for driver distraction, if achieved as expected. However, even in human-robot interaction we could not achieve 100% certainty and accurate results [60]. Moreover, it requires additional cameras and continuous video streaming, in other words very

high processing power. Also, keeping in mind the continually changing environment and bumpy roads, we can never be sure to capture the video.

Another study combines speech and tangible interfaces (turn and push dial) to achieve multimodality in interacting with vehicles [61]. As per the results, there was not much significant difference between speech only and multimodal system.

CHAPTER 8

DESIGN

Navigation Model

A Navigation Model provides a high level overview of a system. It helps users visualize how different parts of any system link with each other at multiple levels. At a simple level, a navigation model helps users understand how one page of an application links to another and how one can navigate through the app. For larger systems, it's important to have good system design as making changes during implementation stages can be time consuming and expensive. This is where navigation models are crucial as they help system designers understand and structure the application better based on the project requirements and give clarity to developers who implement the system.

The model has three layers of navigation primarily:

- a. **Screen to Screen:** This is the most basic layer which shows navigation between two screens. Here is a simple example which shows navigation from a home screen to various parts of the MMIS application like Navigation, Media and Phone.

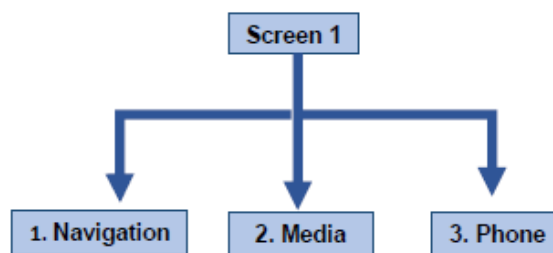


Figure 17: Navigation model: Screen to screen

- b. **Multiple Facets:** This layer is usually used for a screen which has multiple interface related actions which does not take the user to another screen. Best example of this layer would be the concepts of tabs. This kind of navigation is extremely useful when an application has a strong primary view with fewer options. This is because tabs are always visible and a user can quickly navigate to the desired part of the screen. Here is an example

which shows navigation tabular navigation between Photos and Video tabs in a Gallery application available on most smartphone: the overlapping horizontal structure indicates navigation within the same screen i.e. multiple facets of the “Home” screen.

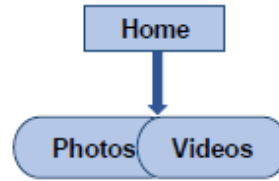


Figure 18: Navigation model: Multiple facets

- c. Superimposed Screens: This layer shows multiple screens which are part of and can be navigated from the same screen but do not act like different facets of the screen. Best way to describe these screens would be that they are superimposed. The best way to illustrate these screens is by placing them in a stack like fashion. An example which shows superimposed screens in the IMMIS application is the Media screen which switches the mode of media. The down arrow indicates that the screens are superimposed.

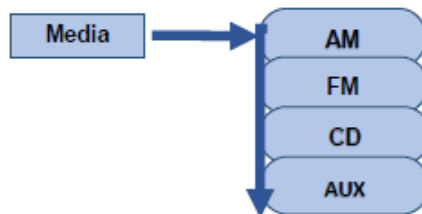


Figure 19: Navigation model: Superimposed screens

Levels of Navigation Model

- a. Level 1: This level gives the high level overview of the navigation of the entire system. It does not concern with details of each screen. A user should be able to see an overview of the pages in an application and how they are connected to each other by looking at the Level 1 model. Best example for this is a sitemap of any website.

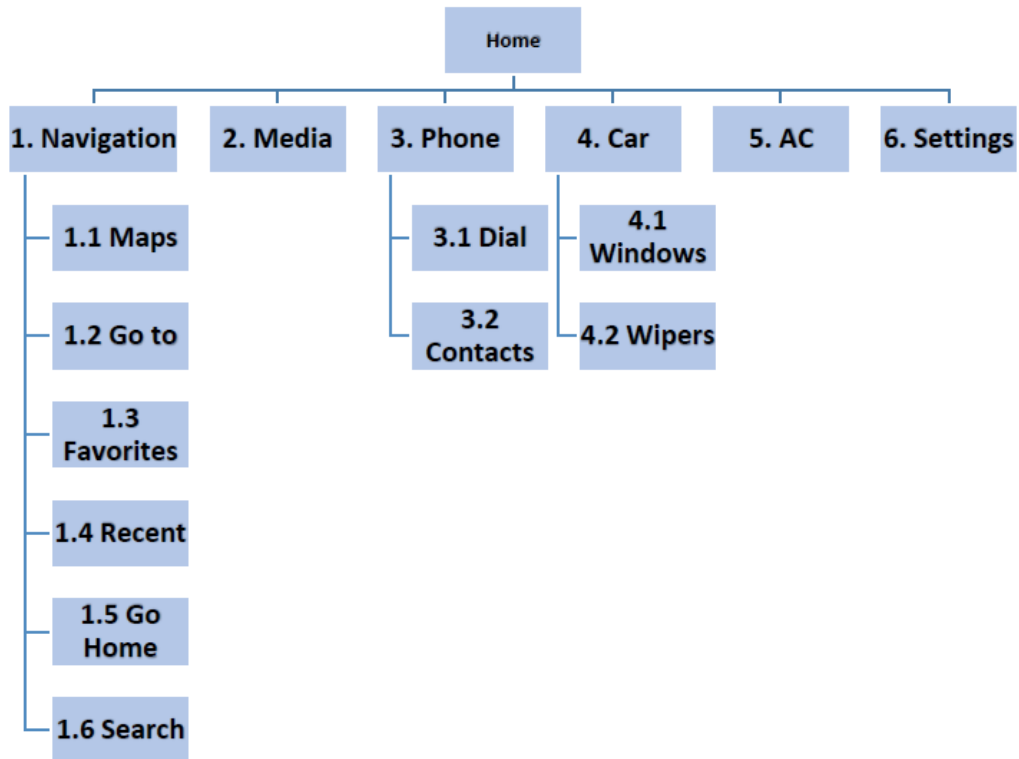


Figure 20: Navigation model: Level 1

- b. Level 2: This level tells the user about the details of individual pages and the actions available to user to navigate to other pages through it. An example of Level 2 would be to show the navigation of any sub screen of an application in detail.

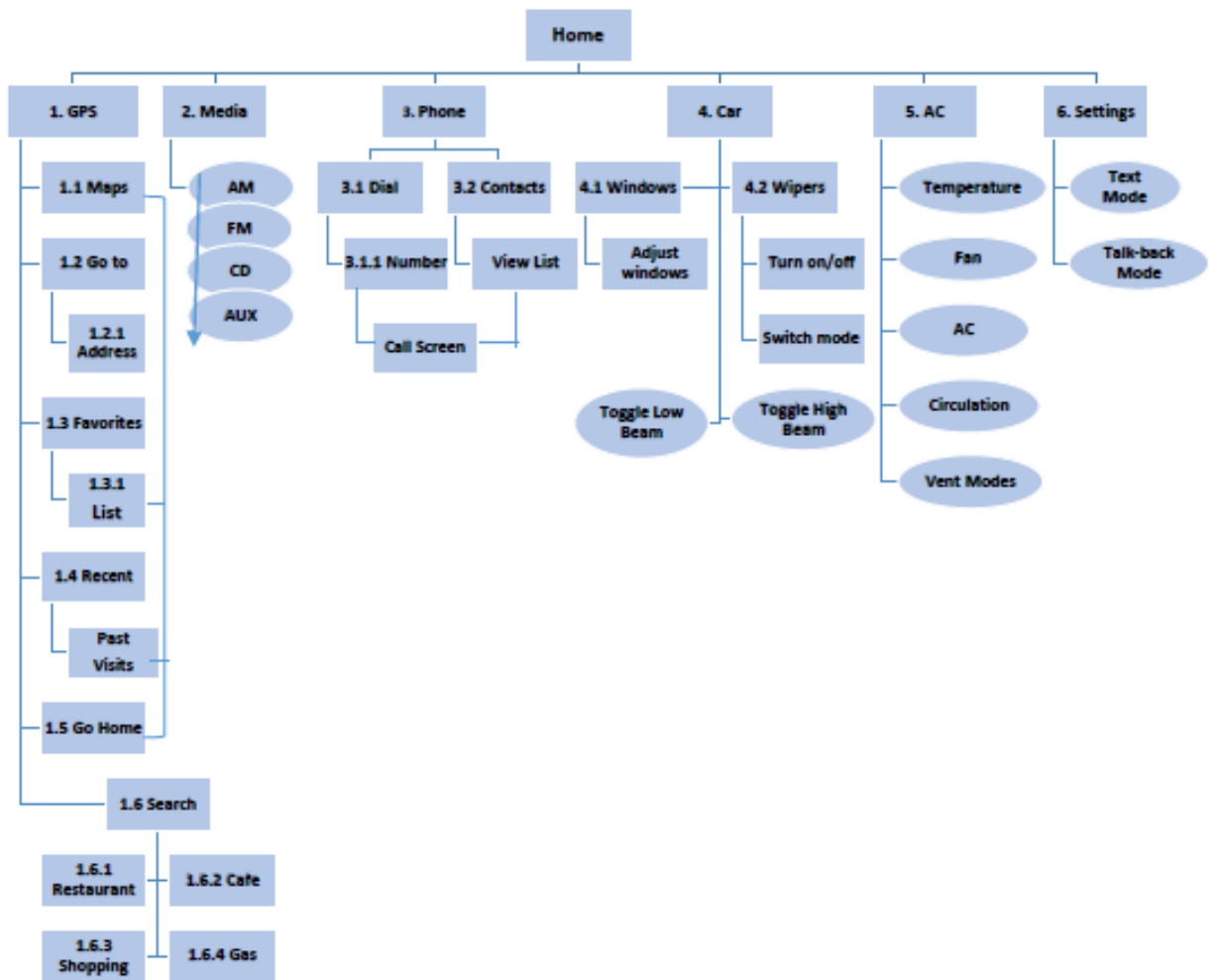


Figure 21: Navigation model: Level 2

- c. Level 3: Third level shows the screens with the components in it and how they link to the other components. This is the most visual and descriptive level and the most difficult to implement at the start of any project. (Figure 22)
- d. The IMMIS application focuses on minimalism, ease of use and driver safety. Therefore it is essential that the driver can navigate to each of the application screens in the least number of steps possible. The navigation model helps us achieve these goals and aids the User Interface designers as well as developers understand the goals of the application better.

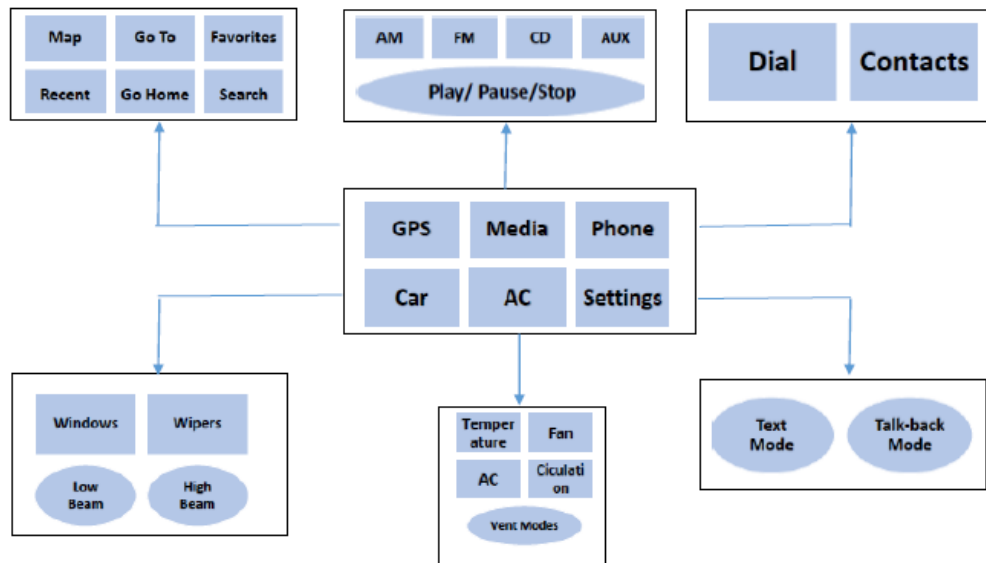


Figure 22: Navigation model: Level 3

User-Interface Design

Design, by definition, is a plan or drawing or an approach produced to show the look or execution of an object before it is made. It is very important to have a good design before starting with actual implementation as Benjamin Franklin said “If you fail to plan, you are planning to fail!” The usability of a product increase when end user’s perception is given the highest priority at the design time and this is termed as User Centered Design. Our main goal is to reduce distraction by minimizing time interacting with the car system. Hence our IMMI User Interface is a hierarchical structure that requires one click per screen and at most 4 clicks to perform an operation correctly. As the human eye perceives images faster than text, we have designed an easy-to-eyes icon layout that requires minimum thought process. The UI we presented is based on a concept called Minimalist design. The term minimalism is also used to describe a trend in design and architecture, wherein the subject is reduced to its necessary elements. Minimalistic design has been highly influenced by Japanese traditional design and architecture. Architect Ludwig Mies van der Rohe adopted the motto "Less is more" to describe his aesthetic tactic of arranging the necessary com-

ponents of a building to create an impression of extreme simplicity—he enlisted every element and detail to serve multiple visual and functional purposes; for example, designing a floor to also serve as the radiator, or a massive fireplace to also house the bathroom. Designer Buckminster Fuller adopted the engineer's goal of "Doing more with less", but his concerns were oriented toward technology and engineering rather than aesthetics. The concept of minimalist architecture is to strip everything down to its essential quality and achieve simplicity. The idea is not completely without ornamentation; but that all parts, details and joinery are considered as reduced to a stage where no one can remove anything further to improve the design.

Following are the few design principles that we have implemented in order to achieve user centered UI design for the infotainment screen [62]:

- **80/20 Rule:** Asserts that approximately 80 percent of the effects generated by 20 percent of the variable in that system. The same rule can be applied to the automotive domain infotainment screen, 80 percent of the usage involves 20 percent of its features. After analyzing different cars and drivers, we have grouped the majorly used features as GPS, music/media, phone connectivity, few car system controls such as wiper and beam, and the air conditioning (Figure 23).
- **Contour Bias:** A tendency to favor objects with contours over objects with sharp angles or points. Instead of keeping sharp edges, we have slightly rounded edges for each icon in our design as shown by pointing an arrow in Figure 23.

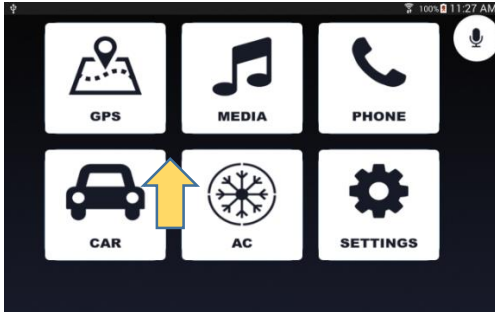


Figure 23: Home Screen with text

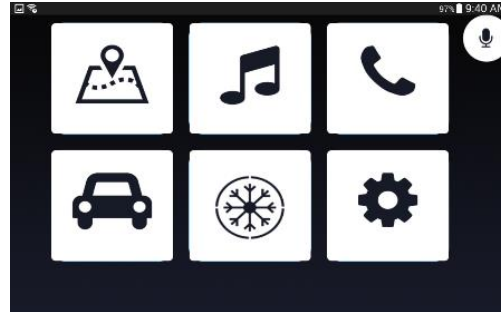


Figure 24: Home screen without text

- **Accessibility:** asserts that designs should be usable by people of diverse abilities without special adaptation or modifications. The four characteristics of accessibility are:
 - Perceptibility: when the design can be perceived by anyone, regardless of sensory abilities.
 - Operability: when anyone can use the design, regardless of physical abilities. The complete application is accessible using both touch and speech.
 - Simplicity: when anyone can understand the design, regardless of experience, literacy, or concentration level. We have tried to keep the graphics and indicative as simple as possible. Also, the text tip feature is an add-on.
 - Forgiveness: when design minimize the occurrence and consequences of errors.

Our UI design provides easily interpretable graphical icons with text as well as without text, depending on the user preference. Also, when the voice-tip mode is switch on then system will speak out the text corresponding to each icon. Figure 23 shows the home screen with text and figure 24 shows the same screen without text.

- **Aesthetic-Usability Effect:** describes a phenomenon in which people perceive more-aesthetic designs as easier to use than less-aesthetic designs due to the lack of acceptance. Example: a circular/triangular/hexagonal shaped infotainment screen might not be accepted by the user. We have tried to keep minimal changes in the UI design that are easy to accept.

- **Alignment:** Elements in a design should be aligned with one or more elements. This creates a sense of unity and cohesion, which contributes to the design's overall aesthetic and perceived stability.

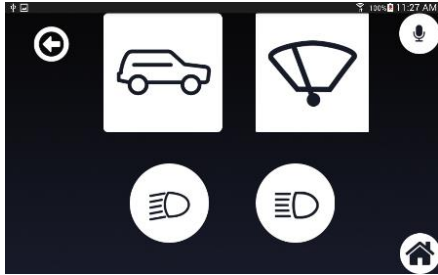


Figure 25: Car systems without text

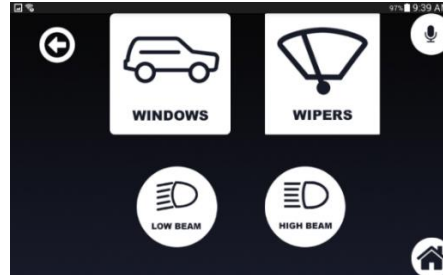


Figure 26: Car systems with text

- **Chunking:** A technique of combining many units of information into a limited number of units or chunks, so that the information is easier to process and remember. We have tried to group similar functional units together such as wipers, windows and lights in car systems as shown in figure 25 and figure 26.
- **Good Continuation:** Elements arranged in a straight line or a smooth curve are perceived as a group and are interpreted as being more related than elements not on the line or curve. (Figure 27)
- **Proximity:** Elements that are close together are perceived to be more related than elements that are farther apart (Figure 27).



Figure 27: Media Player

The media player modes AM, FM and CD are kept in one line whereas the action buttons play, pause and stop represent a different group. The different stations/tracks represented by numbers is another group. (Figure 27)

- **Affordance:** A property in which the physical characteristics of an object or environment influence its function. Round wheels are better suited than square wheels for rolling. This does not say that square wheels cannot rotate but the physical characteristics of round wheels is better suited for the rolling function. The compliance of affordance has been applied throughout the application design. The graphics of icons clearly signify the corresponding function. For example in figure 27, it is quite evident that volume is a seek bar, higher the bar higher is the volume whereas other icons on the screen are buttons.
- **Modularity:** A method of managing system complexity that involves dividing large systems into multiple, smaller self-contained systems. Complying with the rule and not keeping all the possible functions/icons on one screen.
- **Consistency:** Systems are more usable and learnable when similar parts are expressed in similar ways. Four kinds of consistency:
 - Aesthetic: Consistency of style and appearance
 - Functional: Consistency of meaning and action
 - Internal: Consistency with other elements in the section
 - External: Consistency with other elements in the environment

The color coding, shape of icons and font of text is consistent throughout all the screens in application.

- **Control:** The level of control provided by a system should be related to the proficiency and experience levels of the people using the system. The main goal of our design lies in this design principle, the multimodality. The different modes provide user with the leverage to turn on/off based on requirement (Figure 28).

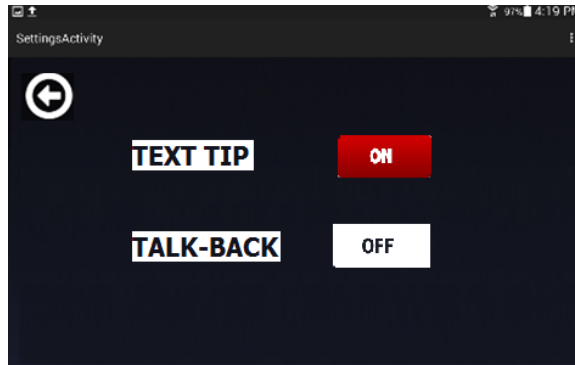


Figure 28: Settings Screen

- **Flexibility-Usability Tradeoff:** As the flexibility of the system increases, the usability of the system decreases. Flexible designs can perform more action but less efficiently. Keeping the number of functions limited in order to reduce distraction as well as increase the efficiency.
- **Golden Ratio:** The ratio within the elements of a form, such as height to width, approximating 0.618. Based on the results of the outreach event, we decided to use 8 inches screen for the center stack.
- **Hick's Law:** The time it takes to make a decision increases as the number of alternatives increases. The system is providing the user with the ability to access only audio applications for entertainment. Features such as video applications and internet browsing are taken off.
- **Hierarchy of Needs:** In order for a design to be successful, it must meet people's basic needs before it can attempt to satisfy higher-level needs.
 - Functionality
 - Reliability
 - Usability
 - Proficiency
 - Creativity

- **Iconic Representation:** The use of pictorial images to improve the recognition and recall of signs and controls. As shown in figure 29, the icon for fan, AC, windshield, face, feet and mix are self-explanatory.

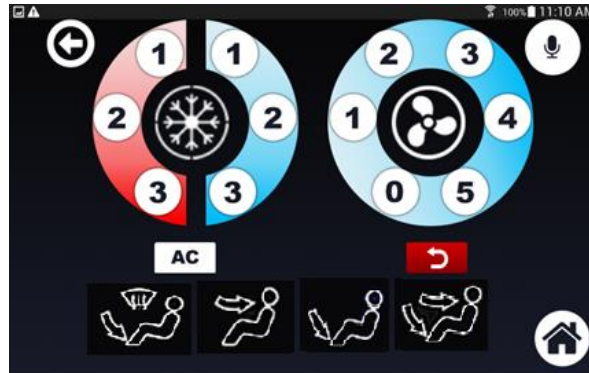


Figure 29: Air Conditioning System

- **Mapping:** A relationship between controls and their movements or effects. Good mapping between controls and their effects results in greater ease of use.
- **Mental Model:** People understand and interact with systems and environments based on mental representations developed from experience. As shown in figure 30, the magnifying glass icon seems familiar and can be easily interpreted as “search”, the star shaped icon gives a clear idea of something being “marked” or “bookmarks” or “favorites”.

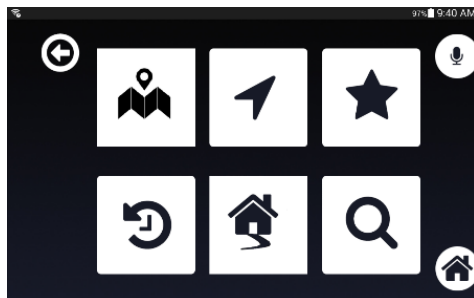


Figure 30: GPS screen without text

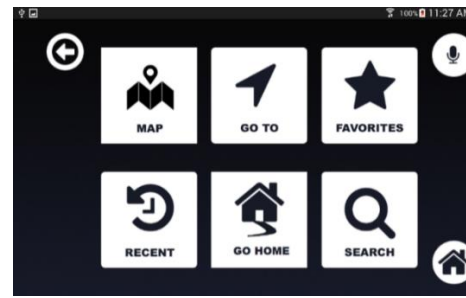
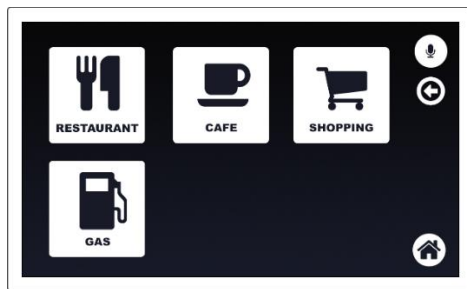


Figure 31: GPS screen with text

- **Orientation sensitivity:** A phenomenon of visual processing in which certain line orientations are more quickly and easily processed and discriminated than other line orientations. Our experiments showed that landscape orientation is preferred over portrait.

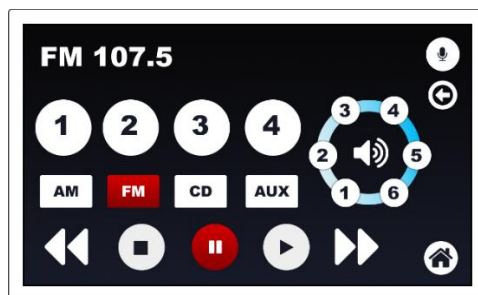
- **Performance Load:** The greater the effort to accomplish a task, the less likely the task will be accomplished successfully. The maximum of number of steps required to perform any task throughout the application is 3-4.
- **Picture Superiority Effect:** Pictures are remembered better than words.
- **Progressive Disclosure:** A strategy of managing information complexity in which only necessary or requested information is displayed at any given time. The text tip mode offers additional information in order to provide the user with some hint. However, the feature can be turned on/off depending on the need.
- **Proportional Density:** The relationship between the elements of a design and the meaning they convey.
- **Prototyping:** The use of simplified and incomplete models of a design to explore ideas, elaborate requirements, refine specifications and test functionality.



Favorites Screen initial version



Favorites screen current version



Media player screen initial version



Media player current version

Figure 32: Prototype images

Figure 32 shows the initial version and the current version of the same screen. Initial version of favorites screen is not utilizing the available space on screen, leaving behind unusual unused space. Also, in the initial version of media player screen, it is difficult to differentiate between the groups for action buttons, media player modes and station/track. The volume is better perceived as a seek bar than numbers.

- **Redundancy:** The use of more elements than necessary to maintain the performance of a system in the event of failure of one or more of the elements. The application is completely accessible using both touch and speech mode.
- **Symmetry:** A property of visual equivalence among elements in a form.
- **Visibility:** The usability of a system is improved when its status and methods of use are clearly visible.

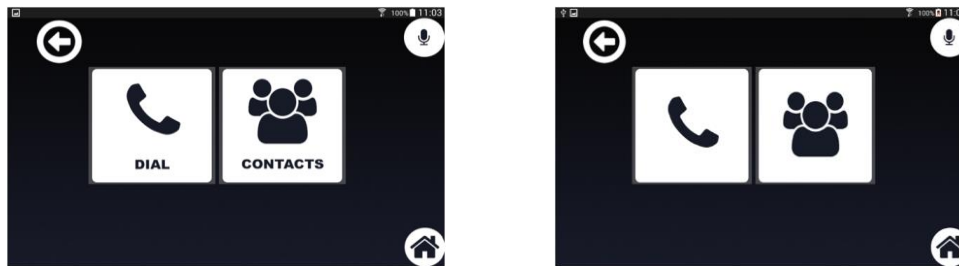


Figure 33: Phone screen with and without text

CHAPTER 9

RELATED PRIOR WORK

Experiment 1

This experiment was conducted as part of Master's Thesis study by Tanvi Jahagirdar [23].

Goal

To evaluate the effect of our minimalist design on driver distraction as well as to measure the effects of icon size and number, screen size and orientation.

An abstract layout of icons of varying sizes, orientation and number of icons while driving, was tested to effectively calculate driver response time.

Setup

The HyperDrive Simulator was used for this experiment. The volunteers were asked to drive on a previously programmed route, with possible driving tasks like left turns at a signal, pedestrians crossing, curved road and following a car. Then the 2 drives – with smaller screen size and larger screen size were monitored closely. The volunteers' reaction time (from the number being said to the driver clicking the number) was also noted for both. For each UI, the driver was asked to click the numbered icon 5 times.

Experiment Data & Analysis

There were 2 types of data obtained – Driver Response Time and Driving Simulation metrics. The response times were closely monitored and noted manually using a stopwatch and excel spreadsheet to note the values. The driving simulation metrics was recorded by the Simulation system by programming it prior to the experiment. All possible metrics which could be used in identifying distraction and its risk factor was separated out in an excel spreadsheet. Since the simulator records every data and action per microsecond, we are unable to display all of the data and only specific parts will be shown.

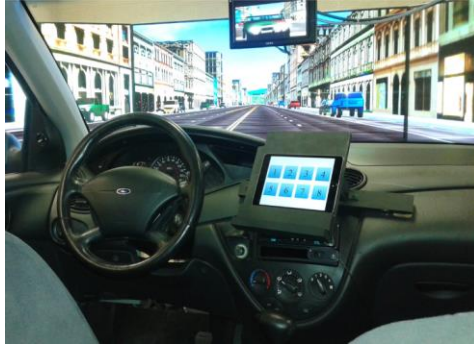


Figure 34: Driver's view of the simulator with 10" tablet displaying 8 icons

Results

Driver Response Time: In general, all response times were below 2 seconds, indicating that this experiment settings can serve as our baseline for the multimodal design. Actual readings ranged from 0.71 seconds to 0.98 seconds for all of the 8 settings. This gave us a strong indication that we can safely design the UI with 8 icons on small screen of 7 inch. A 10 inch screen will not have major improvement if we stayed within our limit of 8 icons. There was no significant difference in the reaction times between portrait and landscape orientation. In the experiment, we interviewed the volunteers to seek their feedback about the screen orientation. Most people preferred Landscape over Portrait orientation. One reasonable explanation is that it is more common as the majority of screen orientations of technology are landscape. For example computers, television, radio and media players are predominantly horizontal. However, more research is needed to test the boundaries of both orientations.

Experiment 2

This experiment was conducted as part of Master's Thesis study by Paresh Nakrani [54].

Goal

To benchmark an abstract screen layout of in-car user interface (UI), to measure the effects of screen size and number of icons on driver distraction and to evaluate the effects of our minimalist design on driver distraction.

Setup

The HyperDrive Simulator was used for this particular experiment. The experiment used two different sizes of android tablets as UIs. The following conventions have been used throughout the experiment:

1) Small Screen: - An android tablet with a 7" (inch) screen is known as "small screen". Small screen has two layouts of icons; first with 24 icons and second with 8 icons.

2) Large Screen: - An android tablet with a 10" inch screen is known as "Large screen". Large Screen has two layouts of icons; first with 24 icons and second with 8 icons.

3) Reaction time (in seconds): - We measure this metric manually and the unit of measurement is in seconds. All four UI's were tested for the driver distraction, and the distraction is measured in terms of reaction time. The reaction time is considered as the amount of time a driver takes his/her eyes off the road while driving to interact with the screen in the center-stack. The Driver is asked to click a certain numbered icon on the screen in the center-stack while driving under the normal conditions. The time taken to click the numbered icon is noted as the reaction time. The greater the reaction time, the higher the distraction. Also, any deviation in those readings without other modifications in the experiment should indicate distraction occurrence. This step is repeated 5 times for each UI screen, and a total of 20 readings are taken per participant, 5 readings per each UI Screen.



Figure 35: Small screen with 24 icons

Experiment Data & Analysis

In this particular experiment, a total of 20 volunteers took part. No personal information was collected about the participants. There were 2 types of data obtained – Driver Reaction Time and Driving Simulation Metrics. The reaction times were closely monitored and noted manually using a stopwatch, and excel spreadsheet was used to note the readings. Driving Simulator Metrics were used to perform statistical analysis of the collected data.

Results

The data collected from the experiment, shows that for the more number of icons (24 icons), the large screen is worse than the small screen and it required more attention, which disproves our null hypothesis that the larger screen is better and less distractive than the smaller screen. SA results also confirms the experiment data.

- There is no statistically significant difference between the small screen and the large screen
- Design with 8 icons was well within the NHTSA's criteria. Total Mean reaction time 1.086 seconds, which nearly the half of the NHTSA's criteria.
- There is an extremely statistically significant difference in the UI with 24 icons and UI with 8 Icons. The mean reaction time for the UI with 24 icons is 1.938120 seconds, which can significantly differ by 0.852 seconds and might cross the limit of 2 seconds. Hence, the UI with 24 icons doesn't meet requirement of NHTSA's guidelines.

CHAPTER 10

EXPERIMENT

An experiment can be defined as a test or a series of test in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in output response [63].

Engineering Based Metrics

Lane Departures

The number of lane departures per unit of time or trial is a very common safety statistic. If drivers are distracted, there are more lane variations. There are two definitions of a lane departure: when “the automobile crossed the white sidelines on the roadway,” and “when the vehicle's tire came into contact with the lane marker.” Thus, there are at least two candidate criteria for a lane departure, (1) the outer edge of the exterior mirror passes over the midline of the lane marking, and (2) the front tire touches the inside edge of the lane marking. The first criterion is the most crash relevant. The second is easier to detect (when using a side-mounted camera). Simple math suggests there is a one to four inch difference between the two criteria [23].

Time-To-Line Crossing

There are actually at least three different ways time-to-line crossing can be defined: (1) as lateral distance divided by lateral velocity, (2) as an expression that includes lateral acceleration, and (3) as the complete trigonometric solution that considers the radius of curvature of the vehicle's path and the radius of curvature of the road. Of the values provided by the three expressions, the first two of which are approximations, and all three can differ considerably. When drivers are distracted, the minimum time-to-line crossing over a time window decreases. Time-to-Line is basically how long it takes the vehicle to reach the lane boundary [23].

Headway

Headway is a measurement of the distance or time between tips of the two vehicles. The more closely one follows a vehicle ahead, the more likely to crash. Of the various types of crashes, rear-end collisions are much more likely when drivers are distracted. Generally, it is the front bumper to front bumper difference.

Acceleration, Velocity, Brake

Previous studies indicate that in distracted driving there is an abrupt change in brake, acceleration and velocity. The driver applies brake or reduces the speed while performing another task and thereafter aligning back to the speed achieved prior to the distraction. This can also indicate the duration of distraction.

Eyes take-off-road Time

We measure this metric manually. The Driver is asked to access a certain application on the screen while driving at normal conditions. In order to measure the visual distraction, the exact duration for which user looked away from the road to see the screen will be recorded [64]. This step will be repeated 4 times for both the test scenarios. And, any deviation in those readings without any other modifications in the experiment should indicate distraction occurrence and larger the time, longer the distraction.

Experiment 1

Statement of the problem:

Is single-mode speech recognition the best solution or is having more modes of interaction better? To get an answer to the above question, we designed our first experiment to compare the command detection hit rate for the same application with single mode speech (we refer to it as the blind Speech Recognition as the user has no idea what the command should be) and dual mode speech with text tip to hint the user to the speech command. Our goal was to evaluate the effect of combining two modes on the system performance.

Null Hypothesis: Dual modal interactive interface has no effect on driver distraction compared to single mode speech recognition.

Alternate Hypothesis: Dual modal interactive interface improves system performance.

Variables

- **Input Variable:** “Text Tip” Mode value- Toggles between “On” and “Off”.
- **Output Response:** Voice Command Hit Rate- The number of attempts to get the command successfully recognized.

Setup

An interface was developed on android platform using Android Studio version 1.2.11 implementing the user centered design principles explained before. The developed application was tested on Android v4.4.2 based Samsung Galaxy Tab4 8.0”. We have designed various prototypes on popular speech recognition technologies such as Android with Google APIs, Python using Google API, IOS with OpenEars, CMU Sphinx using java, Microsoft Windows Speech Recognition using C Sharp etc. Finally, we decided to choose Dragon SDK because of its high accuracy rate in detecting the commands and less processing power requirements.

Dragon SDK version 1.4.11 download was available on the nuance developers’ website <https://dragonmobile.nuancemobiledeveloper.com/public/index.php?task=prodDev#download>.

However in order to get the SDK activated, we had to request the product key with a validity of 90 days from Nuance website.

Steps of Execution

The multimodal application was deployed on a Samsung galaxy tab 4 which was given to the user. Although, user had the leverage to access the application either by touch or by using voice commands, yet for the experimental purpose we restricted the access using voice commands only. Total number of volunteers was 13.

Two scenarios were tested on each participant:

- **Unimodal:** The “Text Tip” mode was switched “off” and user had to access the features or navigate through screens of application using voice commands only.
- **Bimodal:** The “Text Tip” mode was switched “on” and user was asked to access the features or navigate through screens of application only using voice commands.

The command hit rate was captured, i.e., whether the correct command was recognized in one go or in two-three chances or couldn't get recognized. For each participant, it was repeated eleven times on different features.

Results

Figure 36 shows the significant difference between the results of two scenarios. The average success rate (2.6) in unimodal system is significantly less than the average success rate (12.1) in bimodal system. Participants felt a lot easier to access the application with the text tip because of the apparent guidance to a limited set of commands and the less cognitive load in interpreting the icon graphics. Please refer Appendix E section for the table of results.

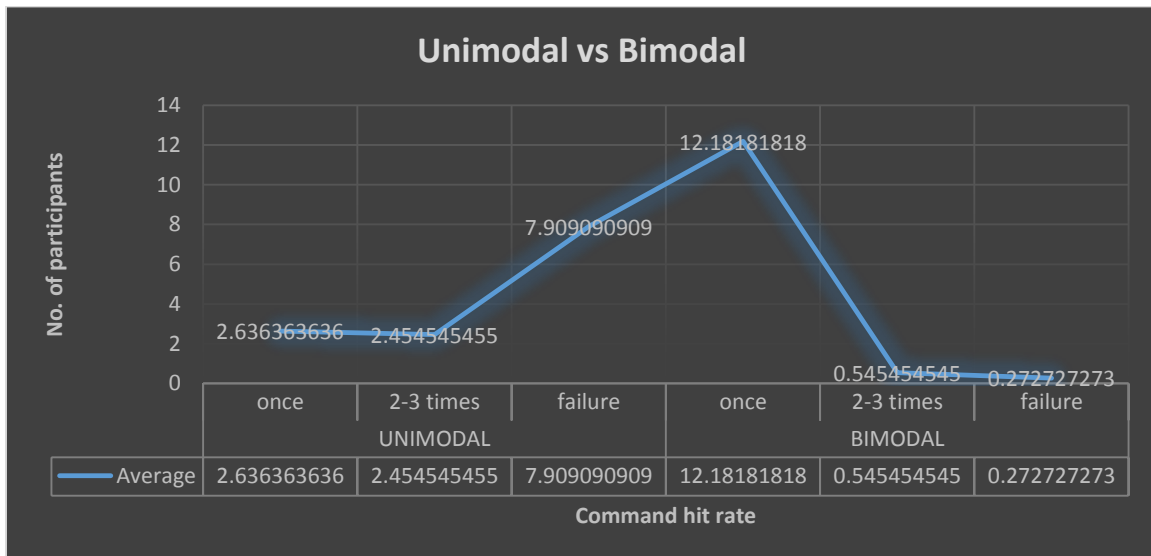


Figure 36: Unimodal vs Bimodal

Experiment 2

Statement of the problem

To compare eyes take-off the road time when using only touch and when using speech mode with text tip mode. Our goal was to evaluate the effect of combining two modes on the driver's performance.

Null Hypothesis: Dual modal interactive interface has no effect on driver distraction compared to single mode speech recognition.

Alternate Hypothesis: Dual modal interactive interface improves driver's performance.

Variables

- **Input Variable:** Speech Command/ touch
- **Output Response:** Eyes take-off the road time

Setup

The HyperDrive Simulator was used for this experiment. It is explained in details later in the Appendix Section. The volunteers were asked to drive on a previously programmed route as shown below in figure 37. Special instructions were given such as not to overtake any car, maintain speed between 40-50 mph and follow all the driving and traffic rules.

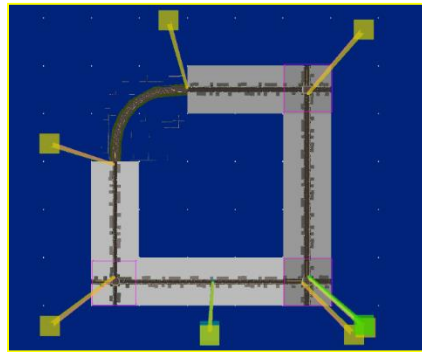


Figure 37: Route overview

At the starting point (figure 38), the user would start from the lane to the right of the centerline. The volunteers were given few minutes to get acquainted with the driving environment. Then the 2 drives – with touch only and speech only were monitored closely. The volunteers' eyes take-off the road time was noted for both the cases.



Figure 38: The starting point

The first complex driving task was a left turn. The driver would have to stop at the intersection and wait for the signal to turn green and then make a turn while following another car also making the same turn. At that point, the drivers were asked to access any center stack application.



Figure 39: The left turn

The curved road was the second complex driving task. At the start of this path, the preceding vehicle is taken out and another car joins the roadway. Drivers were asked to access another application while driving on a curved road.



Figure 40: The curved road

Pedestrian crossing was the last task along with a command given to access the application. Here to capture the eyes take-off the road time more effectively, we had pedestrians crossing the road suddenly. The driver then proceeded forward and reached the goal.

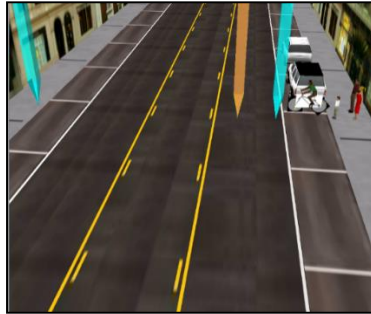


Figure 41: Pedestrian crossing

The same android application which was used in experiment 1 was deployed on an 8 inches android based tablet and was used as the infotainment screen for car. UI design screens can be referred in section.

Special Considerations:

- Due to the noise of simulator, we used a microphone for this experiment (Figure 42).
- Since, there was no dedicated button or activation command for the speech recognition system, we manually clicked the microphone icon present on screen before asking the user to speak.

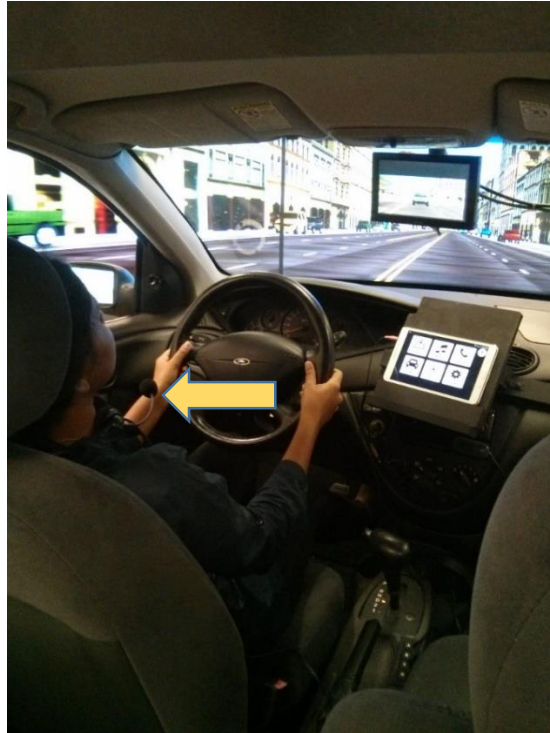


Figure 42: The arrow pointing towards the microphone used for speech recognition

Steps of Execution

- The multimodal application was deployed on a Samsung galaxy tab 4 which was used as the infotainment screen for the driving simulator. Total number of volunteers was 13. At different intervals, users were given few commands such as:
 - You might want to listen to the music.
 - Roll up/down the window
 - You might need to use navigation/GPS system.
 - Look for nearby restaurants or gas stations.
 - You might want to make a call.

Two scenarios were tested on each participant:

- **Single mode only touch:** The “Text Tip” mode was switched “on” and user had to access the features or navigate through screens of application using touch only.

- **Dual mode Speech with Text Tip:** The “Text Tip” mode was switched “on” and user was asked to access the features or navigate through screens of application only using voice commands.

Experiment Data & Analysis

The experiment data shows that there is a significant difference between the two scenarios, yet all of them fall within the NHTSA norms. This proves the effectiveness of our minimalist design and also the comparison of eyes-off-road time proves the multimodal system. That is, each screen should take less than 2 seconds and the cumulative task time should be no more than 2 x 6 screens = 12 seconds. The data thus proves our design causes minimal driver distraction, disproving our null hypothesis, with distractive task time as low as 0.41 seconds per screen to 0.98 seconds. The average eyes-off-road time for multimodal system was 0.74 as compared to 0.89 for unimodal system.

Figure 43 shows the comparison of lane position values with distraction and without distraction when using touch mode and when using speech with text mode. The readings were captured by the driving simulator keeping track of microseconds. The deviation from the center of the lane was recorded in this sample. The graphs are more or less similar.

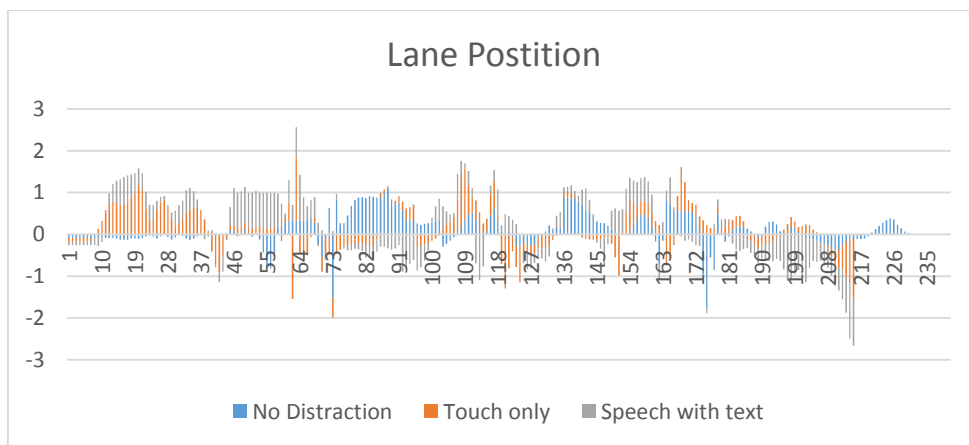


Figure 43: Lane Deviation

The eyes-off-road times were closely monitored and noted manually using a stopwatch and excel spreadsheet to note the values. Rest of the data and statistical analysis is included in the Appendix F.

Results

Figure 44 shows the significant difference between the results of two scenarios. For majority of the scenarios, the time was less for speech with text-tip as compared to the touch mode. However, there were two exceptions where the results were reversed.

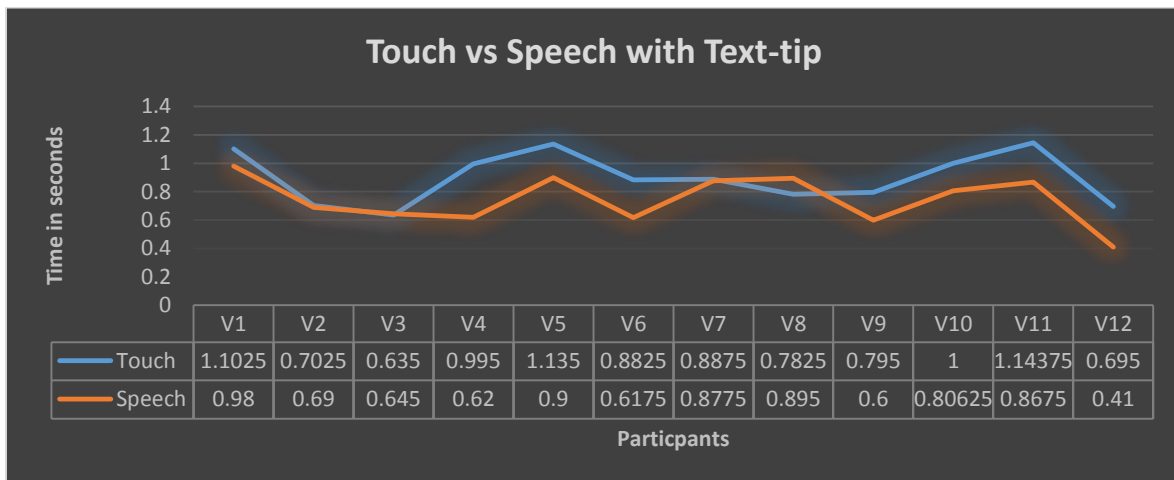


Figure 44: Eyes take-off-road time: Touch vs Speech with text

CHAPTER 11

THE OUTREACH EVENT

Twenty two high school students were brought to the driving simulator lab (SIM 107) for about 2 hours to gain knowledge about the driver distraction and understand the possible consequences of distraction. The event was organized under the supervision of Dr. Ashraf Gaffar and Dr. Ashish Amresh. The students were really excited to see the simulator and few of them seemed very interested in gaining further knowledge about the research. We started with introduction followed by a query session and then each student was given a chance to drive the simulator for approx. 7min.

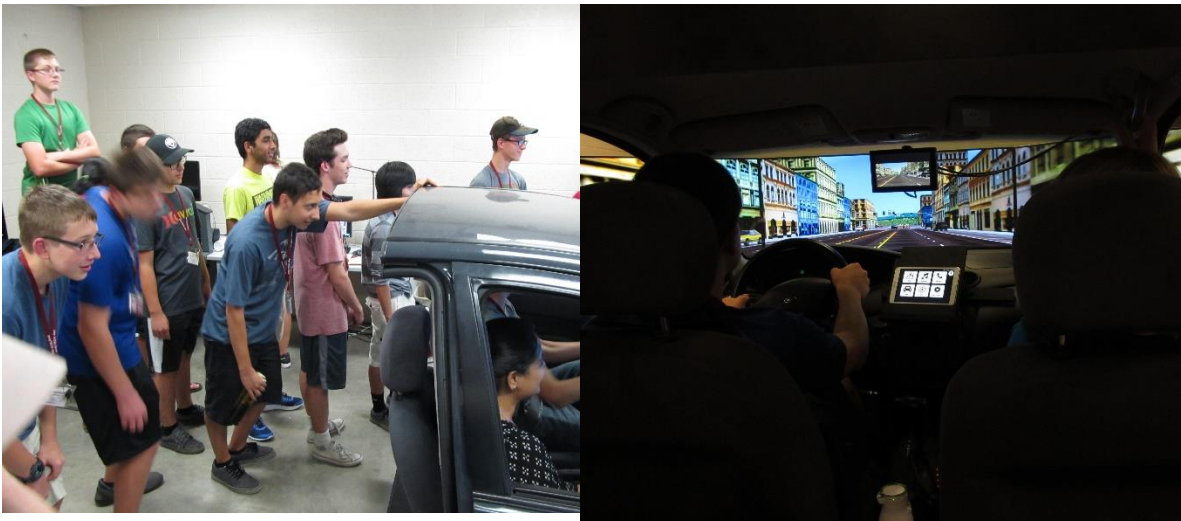


Figure 45: The outreach event

Introduction provided to the students

Our research is focused on looking out the factors that cause driver distraction and measures to reduce the same. Today, the car is nothing but driving a computer on wheels. We have a whole bunch of distractive features such as calling, texting, social media, web browsing etc. and whether required or not driver tends to use them. These features are increasing distraction. A distracted driver is not only keeping his/her own life at risk but also the life of other innocent souls as well.

Query Session

Question to the students: Why do you think the distraction will increase in near future?

Answer: Because the number of features in cockpit are increasing and soon it will be nothing less but driving a computer with wheels.

Question by one of the students: Does the simulator have 3D effect? Can we see the cars flying when met with an accident?

Answer: No, this simulator is designed with an educational purpose to measure distraction not to show the graphics.

Question by Dr. Amresh to us: Instead of providing a solution to minimize driver distraction, why can't we provide the training to drive with distraction?

Answer: We are not expecting that the driver would not use cell phones/center stack while driving. However, we can also not expect to train everyone for multitasking. A solution in mid-way of the two is required indeed.



Figure 46: Students filling out the questionnaire at the outreach event

Result

A questionnaire (Appendix D) was prepared and students were asked to fill that out. Results showed that most of the students were in the age group of 14-16 years with a few exceptions. There were 20 male and 2 female students. We identified three groups of students: ones with driving license, second group for students without the license but quite familiar with driving and the third group with no driving experience at all. The first group with 25% of the students were the first ones to drive on the simulator. Second group contained the majority, i.e., 57% of the students followed by 18% in the third group. All of them were aware of the legal age to get driver's license and were quite well informed about the dos and don'ts of driving. All of them said a driver must not use cell phone or texting while driving. However, 36% of the students said music player can be accessed while driving. Their perception of distraction implied texting/calling/navigation system while driving, temptation or use of technology and social media, boredom, lack of sleep, other people etc. Lastly, we showed a few tablets with size range from 7-10 inches and asked for the student's preference to determine the aspect ratio. 50% of the students preferred 7 inches tablets as compared to 20 % preference for 8 inches and 15% for 10 inches while 15 % gave no preference at all.



Figure 47: Students enjoying while driving the simulator

Observation

It was an overwhelming experience to see high school students, the youth being aware of the dos and don'ts of driving. We observed that distraction is not just because of the external factors but is also dependent on the person's state of mind and concentration power. During the event, we realized there were few students who were getting way too distracted by the other students whereas few of them didn't really bother. Students with good confidence in their driving did a much better job than others.

CHAPTER 12

CONCLUSION

After the careful analysis of factors affecting driver distraction and human-human multimodal interaction, we claim that working towards perfection in single mode speech oriented interaction is not the best solution to driver distraction. We propose IMMI to be a prospective approach. Following are the conclusion drawn based on the master's thesis research work:

- **Multiple modes are better:** The result of experiments evidently proved the combination of four modes (touch, speech, graphics, and text tip) has proven effective. The offered "text tip" provided the user with a better idea of command to be used to access the application.
- **Minimalist Design** with number of icons per screen within the range 8-10 max proved to be an effective design. The average of eyes-off-road time was less than 1sec for all the participants. The time required to look for an icon on a screen either clutter with lots of icons or small sized icons can lead to a high eyes-off-road time.
- **Keyword based speech recognition** is much more effective than having a huge spectrum of natural language speaking. The more options are there, the more are the number of changes to get it misdiagnosed. This improves the command detection and accuracy rate.
- **Eyes-off-road time** for single mode is greater than that for dual modal system. Apparently, this was the first time for all the volunteers to use the system, so they had to look towards the screen to read out the offered "text tip". However, within two-three practices, driver will get acquainted with system and there would not be a need to look through.

CHAPTER 13

FUTURE WORK

This above work was conducted on a limited number of participants. More test cases and more participants should definitely show more accurate results. As explained in section, the voice-tip mode and the learning mode could not be implemented as a part of this thesis and hence, will be taken care of in future work. Few more experiments can be conducted combining the different mode and comparing the result. As a part of current study, we just measured visual and manual distraction, but yet to measure the cognitive distraction. There is also a potential of identifying more such automotive context specific modes. The careful combination of the modes as well as the inclusion of artificial intelligence algorithms to detect failed recognition and improve system learnability should provide a much richer interactive environment between the driver and the car for more natural and less distractive user experience. ADAS system can be combined with IMMI to make it more efficient.

There is a need to check the difference between normal and emergency conditions. The reactions and situations are different in real scenarios. Also, the microphone was kept very near to the mouth. The position we placed our screen was right next to the steering wheel. However there could be further research concerning to this case.

REFERENCES

- [1] B. of T. S. US Department of Transportation, "The Changing Face of Transportation," *Wash. DC2000*.
- [2] J. D. Lee, "Technology and teen drivers," *J. Safety Res.*, vol. 38, no. 2, pp. 203–213, 2007.
- [3] "General yearly statistics for motor vehicle crashes by Insurance Institute for Highway Safety Highway Loss Data Institute."
- [4] D. I. D. US Census Bureau, "Population Estimates." [Online]. Available: <http://www.census.gov/popest/data/datasets.html>. [Accessed: 14-Jun-2015].
- [5] B. of T. S. U.S. Department of Transportation, "Transportation Statistics Annual Report 2013," *Wash. DC 2014*.
- [6] J. R. Treat, N. S. Tumbas, S. T. McDonald, D. Shinar, R. D. Hume, R. E. Mayer, R. L. Stansifer, and N. J. Castellan, "Tri-level study of the causes of traffic accidents: final report. Executive summary.," May 1979.
- [7] J. D. Lee, "Driving Safety," *Rev. Hum. Factors Ergon.*, vol. 1, no. 1, pp. 172–218, 2005.
- [8] Y. Liang, "Detecting driver distraction," *Theses Diss.*, Jan. 2009.
- [9] "Key Facts & Statistics for Distracted Driving by US Government."
- [10] NHTSA, "U.S. DOT Releases Guidelines to Minimize In-Vehicle Distractions," Apr. 2013
- [11] American Automobile Association, "Distracted-Driving."
- [12] D. L. Strayer, J. M. Cooper, J. Turrill, J. Coleman, N. Medeiros-Ward, and F. Biondi, "Measuring cognitive distraction in the automobile," 2013.
- [13] J. M. Cooper, N. Medeiros-Ward, and D. L. Strayer, "The Impact of Eye Movements and Cognitive Workload on Lateral Position Variability in Driving," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 55, no. 5, pp. 1001–1014, Oct. 2013.
- [14] J. L. Harbluk, Y. I. Noy, and M. Eizenman, "The impact of cognitive distraction on driver visual behaviour and vehicle control," 2002.
- [15] *The Psychology of Learning and Motivation: Advances in Research and Theory*. Academic Press, 2011.
- [16] F. J. Lee and N. A. Taatgen, "Multitasking as skill acquisition," in *Proceedings of the twenty-fourth annual conference of the cognitive science society*, 2002, pp. 572–577.
- [17] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learn. Instr.*, vol. 4, no. 4, pp. 295–312, 1994.
- [18] National Highway Traffic Safety Administration, "Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices."
- [19] NHTSA, "NHTSA Survey Finds 660,000 Drivers Using Cell Phones or Manipulating Electronic Devices While Driving At Any Given Daylight Moment," Apr. 2013.
- [20] "Distracted Driving." [Online]. Available: <https://www.aaafoundation.org/distracted-driving>. [Accessed: 15-Jun-2015].
- [21] N. H. T. S. Administration and others, "Traffic Safety Facts: 2012 Data," *Wash. DC US Dep. Transp. Natl. Highw. Traffic Saf. Adm. Online*, 2014.

- [22] K. Young, M. Regan, and M. Hammer, "Driver distraction: A review of the literature," *Distracted Driv.*, pp. 379–405, 2007.
- [23] T. Jahagirdar, A. Gaffar, A. Ghazarian, R. Gray, and Arizona State University, "Modeling and Measuring Cognitive Load to Reduce Driver Distraction in Smart Cars," in *ASU Electronic Dissertations and Theses*, Arizona State University, 2015.
- [24] L. Jin, Q. Niu, H. Hou, H. Xian, Y. Wang, and D. Shi, "Driver Cognitive Distraction Detection Using Driving Performance Measures," *Discrete Dyn. Nat. Soc.*, vol. 2012, p. e432634, Nov. 2012.
- [25] "Motoring with microprocessors," *Embedded*. [Online]. Available: <http://www.embedded.com/electronics-blogs/significant-bits/4024611/Motoring-with-microprocessors>. [Accessed: 02-Jul-2015].
- [26] F. J. Mammano and J. J.R. Bishop, "Status of IVHS technical developments in the United States," in *Vehicular Technology Conference, 1992, IEEE 42nd*, 1992, pp. 85–88 vol.1.
- [27] A. Talati Vaishakhi and V. Talati Ashish, "Innovative Transportation Technique: A Need for Urban Traffic Control, Regulation and Management."
- [28] "Advanced Avionics Handbook." [Online]. Available: http://www.faa.gov/regulations_policies/handbooks_manuals/aviation/advanced_avionics_handbook/. [Accessed: 13-Jul-2015].
- [29] D. A. Norman, "The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 327, no. 1241, pp. 585–593, 1990.
- [30] D. Javaux and V. De Keyser, "The cognitive complexity of pilot-mode interaction," 1998.
- [31] C. E. Billings, *Aviation automation: The search for a human-centered approach*. 1997.
- [32] J. Rushby, "Using model checking to help discover mode confusions and other automation surprises," *Reliab. Eng. Syst. Saf.*, vol. 75, no. 2, pp. 167–177, 2002.
- [33] N. D. Sarter and D. and Woods, "How in the World did I Ever Get into That Mode?": Mode Error and Awareness in Supervisory Control," *Human Factors* 37, pp. 5–19.
- [34] A. Degani and M. Heymann, "Pilot-autopilot interaction: A formal perspective," *Abbott A11*, pp. 157–168, 2000.
- [35] P. Mellor, "CAD: computer-aided disaster," *HIGH INTEGR SYST*, vol. 1, no. 2, pp. 101–156, 1994.
- [36] "Radio Technical Commission for Aeronautics," *Wikipedia, the free encyclopedia*. 31-May-2015.
- [37] "What are drones?," *Drone Wars UK* .
- [38] "UAV."
- [39] "Autonomous car," *Wikipedia, the free encyclopedia*. 05-Jul-2015.
- [40] "Google driverless car," *Wikipedia, the free encyclopedia*. 08-Jul-2015.
- [41] C. Urmson, "The self-driving car logs more miles on new wheels," *Google Off. Blog*, 2012.
- [42] E. Guizzo, "How google's self-driving car works," *IEEE Spectr. Online Oct.*, vol. 18, 2011.

- [43] D. J. Fagnant and K. Kockelman, "PREPARING A NATION FOR AUTONOMOUS VEHICLES: 1 OPPORTUNITIES, BARRIERS AND POLICY RECOMMENDATIONS FOR 2 CAPITALIZING ON SELF-DRIVEN VEHICLES 3," *Transp. Res.*, vol. 20, 2014.
- [44] M. M. Azmat and C. Schuhmayer, "Self Driving Cars," 2015.
- [45] M. L. Cummings and J. C. Ryan, "Shared Authority Concerns in Automated Driving Applications," 2014.
- [46] B.-H. Juang and L. R. Rabiner, "Automatic speech recognition—A brief history of the technology development," *Encycl. Lang. Linguist.*, pp. 1–24, 2005.
- [47] "Speech recognition," *Wikipedia, the free encyclopedia*. 04-Jun-2015.
- [48] J. Baker, L. Deng, J. Glass, S. Khudanpur, C.-H. Lee, N. Morgan, and D. O'Shaughnessy, "Developments and directions in speech recognition and understanding, Part 1 [DSP Education]," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 75–80, May 2009.
- [49] "Hype cycle," *Wikipedia, the free encyclopedia*. 07-May-2015.
- [50] "Hype Cycle for Human-Computer Interaction, 2014." [Online]. Available: <https://www.gartner.com/doc/2822817/hype-cycle-humancomputer-interaction->. [Accessed: 11-Jun-2015].
- [51] "Voice recognition installed in new cars 2019 | Forecast," *Statista*. [Online]. Available: <http://www.statista.com/statistics/261685/voice-recognition-in-new-cars/>. [Accessed: 11-Jun-2015].
- [52] D. L. Strayer, J. Turrill, J. R. Coleman, E. V. Ortiz, and J. M. Cooper, "Measuring Cognitive Distraction in the Automobile II: Assessing In-Vehicle Voice-Based Interactive Technologies," Oct. 2014.
- [53] M. C. McCallum, J. L. Campbell, J. B. Richman, J. L. Brown, and E. Wiese, "Speech Recognition and In-Vehicle Telematics Devices: Potential Reductions in Driver Distraction," *Int. J. Speech Technol.*, vol. 7, no. 1, pp. 25–33, Jan. 2004.
- [54] P. K. Nakrani, A. Gaffar, S. Sohoni, A. Ghazarian, and Arizona State University, "Smart Car Technologies: A Comprehensive Study of the State of the Art with Analysis and Trends," in *ASU Electronic Dissertations and Theses*, Arizona State University, 2015.
- [55] S. P. A. Ringland and F. J. Scahill, "Multimodality — The Future of the Wireless User Interface," *BT Technol. J.*, vol. 21, no. 3, pp. 181–191, Jul. 2003.
- [56] T. Stivers and J. Sidnell, "Introduction: Multimodal interaction," *Semiotica*, vol. 2005, no. 156, pp. 1–20, Dec. 2005.
- [57] A. Jaimes and N. Sebe, "Multimodal human computer interaction: A survey," in *Computer vision in human-computer interaction*, Springer, 2005, pp. 1–15.
- [58] M. Turk, "Multimodal human-computer interaction," in *Real-time vision for human-computer interaction*, Springer, 2005, pp. 269–283.
- [59] T. M. Sezgin, I. Davies, and P. Robinson, "Multimodal inference for driver-vehicle interaction," in *Proceedings of the 2009 international conference on Multimodal interfaces*, 2009, pp. 193–198.
- [60] G. Littlewort, M. S. Bartlett, I. R. Fasel, J. Chenu, T. Kanda, H. Ishiguro, and J. R. Movellan, "Towards Social Robots: Automatic Evaluation of Human-robot Interaction by Face Detection and Expression Classification.," in *NIPS*, 2003.

- [61] S. Castronovo, A. Mahr, M. Pentcheva, and C. Müller, "Multimodal dialog in the car: combining speech and turn-and-push dial to control comfort functions," in *Eleventh Annual Conference of the International Speech Communication Association*, 2010.
- [62] W. Lidwell, K. Holden, and J. Butler, *Universal Principles of Design, Revised and Updated: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design*. Rockport Publishers, 2010.
- [63] D. C. Montgomery, *Design and Analysis of Experiments*. John Wiley & Sons, 2008.
- [64] P. Burns, J. Harbluk, J. P. Foley, and L. Angell, "The importance of task duration and related measures in assessing the distraction potential of in-vehicle tasks," in *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2010, pp. 12–19.

APPENDIX A

TABLE USED FOR THE POPULATION CHART

Year	Death	Death Rate	Crash	Crash Rate	Motor vehicles	Motor Vehicle Rate	Population	Population Rate
1975	44,525		39,161		55,534		215,973,000	
1977	47,878	5.173	42,211	6.199	60,516	7.902	220,239,000	1.011
1979	51,093	1.514	45,223	1.778	64,762	0.963	225,055,000	1.110
1981	49,301	-3.504	44,000	-2.835	62,699	-1.238	229,966,000	0.984
1982	43,945	-10.864	39,092	-11.155	56,455	-9.959	232,188,000	0.966
1983	42,589	-3.086	37,976	-2.855	55,106	-2.390	234,307,000	0.913
1984	44,257	3.917	39,631	4.358	57,972	5.201	236,348,000	0.871
1985	43,825	-0.976	39,196	-1.098	58,272	0.517	238,466,000	0.896
1987	46,390	0.657	41,438	0.847	61,836	1.717	242,804,000	0.895
1989	45,582	-3.196	40,741	-3.297	60,870	-2.923	247,342,000	0.947
1991	41,508	-6.931	36,937	-7.277	54,795	-7.584	253,493,000	1.344
1993	40,150	2.293	35,780	2.398	53,777	2.968	260,255,000	1.308
1994	40,716	1.410	36,254	1.325	54,911	2.109	263,436,000	1.222
1995	41,817	2.704	37,241	2.722	56,524	2.937	266,557,000	1.185
1996	42,065	0.593	37,494	0.679	57,347	1.456	269,667,000	1.167
1997	42,013	-0.124	37,324	-0.453	57,060	-0.500	272,912,000	1.203
1998	41,501	-1.219	37,107	-0.581	56,922	-0.242	276,115,000	1.174
1999	41,717	0.520	37,140	0.089	56,820	-0.179	279,295,000	1.152
2000	41,945	0.547	37,526	1.039	57,594	1.362	282,384,579	1.106
2001	42,196	0.598	37,862	0.895	57,918	0.563	285,309,019	1.036
2002	43,005	1.917	38,491	1.661	58,426	0.877	288,104,818	0.980
2003	42,884	-0.281	38,477	-0.036	58,877	0.772	290,819,634	0.942
2004	42,836	-0.112	38,444	-0.086	58,729	-0.251	293,463,185	0.909
2005	43,510	1.573	39,252	2.102	59,495	1.304	296,186,216	0.928
2006	42,708	-1.843	38,648	-1.539	58,094	-2.355	298,995,825	0.949
2007	41,259	-3.393	37,435	-3.139	56,253	-3.169	302,003,917	1.006
2008	37,423	-9.297	34,172	-8.716	50,660	-9.943	304,797,761	0.925
2009	33,883	-9.459	30,862	-9.686	45,540	-10.107	307,439,406	0.867

Table 4: Crash, death, vehicle, population growth rate

APPENDIX B
DRIVING SIMULATOR

The experiment was conducted in a DriveSafety research simulator DS-600s (Figure 48). The DS-600c is a fully integrated, high performance, high fidelity driving simulation system which includes multi-channel audio/visual systems, a minimum 180° wraparound display, full-width automobile cab (Ford Focus) including windshield, driver and passenger seats, center console, dash and instrumentation, and real-time vehicle motion simulation. It renders visual imagery at 60 frames per second on a sophisticated out-the-window visual display with horizontal field-of-view. It also includes three independently configurable rear view mirrors.



Figure 48: Driving simulator

Figure 48 shows the driving simulator computer used for designing and executing the simulation. All driving scenarios were created using DriveSafety HyperDrive Authoring Suite version 1.9.39. HyperDrive is an integrated, Windows-based software package that lets you develop driving simulation content for your simulator. With HyperDrive's point-and-click, drag-and-drop interface, even non-technical users can design, build, execute, and analyze driving scenarios. These vehicles obey traffic laws, signs and signal devices, and interact realistically with other vehicles based on human behavior and real-time physics-based vehicle dynamics. If specific behavior is desired, vehicles can be given script commands through the use of triggers, timers, paths, routes and other tools. We can control traffic signals, ambient traffic, scripted traffic, roadway friction, weather conditions, etc. Through the use of triggers, virtually any scenario can be designed. The rear view mirrors also have small tablets displaying the rear end of car to keep track of blind spot and lane.

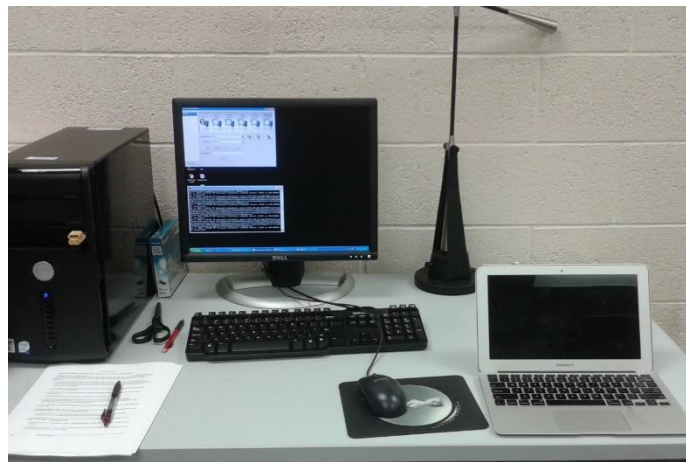


Figure 49: Driving simulator Computer

Figure 50 shows the center stack screen holder to support different types of screens ranging from size 7" to 10" portrait as well as landscape orientation.

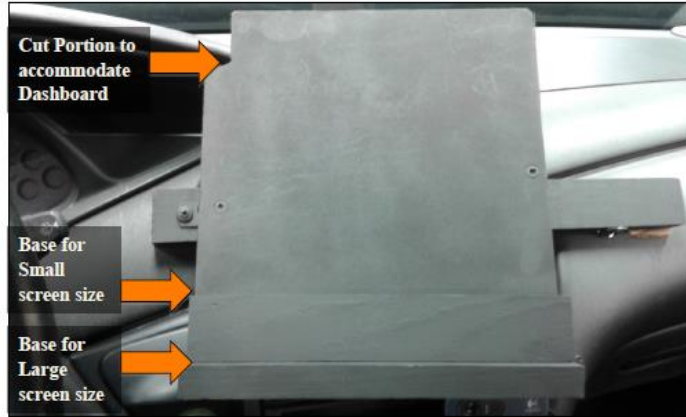


Figure 50: Center stack screen holder

Since, we wanted the screen to be slightly angled to the driver and at a height, which does not result in total visual distraction. Thus it was affixed at a height that could be seen from the corner of our eye, without losing visual on the road. To angle it towards the driver, the support between the back and front was cut of different lengths. And a base for the screens to rest on was fixed at the bottom.

APPENDIX C
STATISTICAL ANALYSIS

Statistics is the study of the collection, analysis, interpretation, presentation, and organization of data. Statistics deals with all aspects of data including the planning of data collection in terms of the design of surveys and experiments. Two main statistical methodologies are used in data analysis:

- **Descriptive** statistics, which summarizes data from a sample using indexes such as the mean or standard deviation, and inferential statistics, which draws conclusions from data that are subject to random variation (e.g., observational errors, sampling variation).
- **Standard** statistical procedure involves the development of a null hypothesis, a general statement or default position that there is no relationship between two quantities.

Rejecting or disproving the null hypothesis is a central task in the modern practice of science, and gives a precise sense in which a claim is capable of being proven false. What statisticians call an alternative hypothesis is simply a hypothesis that contradicts the null hypothesis.

Statistical analysis is a component of data analytics. In the context of business intelligence (BI), statistical analysis involves collecting and scrutinizing every single data sample in a set of items from which samples can be drawn. Statistical analysis is fundamental to all experiments that use statistics as a research methodology. Most experiments in social sciences and many important experiments in natural science and engineering need statistical analysis. Statistical analysis is also a very useful tool to get approximate solutions when the actual process is highly complex or unknown in its true form. Example: The study of turbulence relies heavily on statistical analysis derived from experiments. Turbulence is highly complex and almost impossible to study at a purely theoretical level. Scientists therefore need to rely on a statistical analysis of turbulence through experiments to confirm theories they propound. In social sciences, statistical analysis is at the heart of most experiments. It is very hard to obtain general theories in these areas that are universally valid. In addition, it is through experiments and surveys that a social scientist is able to confirm his theory.

Statistical analysis can be broken down into five discrete steps, as follows:

1. Describe the nature of the data to be analyzed.
2. Explore the relation of the data to the underlying population.
3. Create a model to summarize understanding of how the data relates to the underlying population.
4. Prove (or disprove) the validity of the model.
5. Employ predictive analytics to run scenarios that will help guide future actions.

The goal of statistical analysis is to identify trends. A retail business, for example, might use statistical analysis to find patterns in unstructured and semi-structured customer data that can be used to create a more positive customer experience and increase sales.

Calculation of the test statistic requires four components:

- The average of the sample (observed average)
- The population average or other known value (expected average)
- The standard deviation (SD) of the sample average
- The number of observations.

With these four pieces of information, we calculate the following statistic, t:

$$t = \frac{\text{observed} - \text{expected}}{\text{SD}_{\text{observed}} \sqrt{(\text{number of observations in sample}) / ((\text{number of observation}) - 1)}}$$

A single sample t-test (or one sample t-test) is used to compare the mean of a single sample of scores to a known or hypothetical population mean. So, for example, it could be used to determine whether the mean diastolic blood pressure of a particular group differs from 85, a value determined by a previous study.

Requirements

- The data is normally distributed
- Scale of measurement should be interval or ratio
- A randomized sample from a defined population

- Null Hypothesis

H0: $M - \mu = 0$, where M is the sample mean and μ is the population or hypothesized mean. As above, the null hypothesis is that there is no difference between the sample mean and the known or hypothesized population mean.

Equation:

$$t = (M - \mu) / \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{(N-1)(N)}}$$

Suppose that you've collected data from two samples of animals treated with different drugs. You've measured an enzyme in each animal's plasma, and the means are different. You want to know whether that difference is due to an effect of the drug – whether the two populations have different means. Observing different sample means is not enough to persuade you to conclude that the populations have different means. It is possible that the populations have the same mean (i.e., that the drugs have no effect on the enzyme you are measuring) and that the difference you observed between sample means occurred only by chance. There is no way you can ever be sure if the difference you observed reflects a true difference or if it simply occurred in the course of random sampling. All you can do is calculate probabilities. The P value is a probability, with a value ranging from zero to one that answers this question (which you probably never thought to ask): In an experiment of this size, if the populations really have the same mean, what is the probability of observing at least as large a difference between sample means as was, in fact, observed?

The confidence interval (CI) of a mean tells you how precisely you have determined the mean. In statistics, the number of degrees of freedom (df) is the number of values in the final calculation of a statistic that are free to vary. The standard deviation (SD) quantifies variability. It is expressed in the same units as the data. The Standard Error of the Mean (SEM) quantifies the precision of the mean. It is a measure of how far your sample mean is likely to be from the true population mean. It is expressed in the same units as the data. For example, you measure weight in a small sample (N=5), and compute the mean. That mean is very unlikely to equal the population mean.

The size of the likely discrepancy depends on the size and variability of the sample. If your sample is small and variable, the sample mean is likely to be quite far from the population mean. If your sample is large and has little scatter, the sample mean will probably be very close to the population mean. Statistical calculations combine sample size and variability (standard deviation) to generate a CI for the population mean. As its name suggests, the CI is a range of values. To interpret the confidence interval of the mean, you must assume that all the values were independently and randomly sampled from a population whose values are distributed according to a Gaussian distribution. If you accept those assumptions, there is a 95% chance that the 95% CI contains the true population mean. In other words, if you generate many 95% CIs from many samples, you can expect the 95% CI to include the true population mean in 95% of the cases, and not to include the population mean value in the other 5%. The unpaired t test compares the means of two unmatched groups, assuming that the values follow a Gaussian distribution. The unpaired t test assumes that the two populations have the same variances (and thus the same standard deviation). The paired t test compares the means of two matched groups, assuming that the distribution of the before-after differences follows a Gaussian distribution. The paired t test assumes that you have sampled your pairs of values from a population of pairs where the difference between pairs follows a Gaussian distribution.

Note that the paired t test, unlike the unpaired t test, does not assume that the two sets of data (before and after, in the typical example) are sampled from populations with equal variances. The pairing should be part of the experimental design and not something you do after collecting data. Prism tests the effectiveness of pairing by calculating the Pearson correlation coefficient, r , and a corresponding P value. If the P value is small, the two groups are significantly correlated. This justifies the use of a paired test. If this P value is large (say larger than 0.05), you should question whether it made sense to use a paired test. Your choice of whether to use a paired test or not should not be based solely on this one P value, but also on the experimental design and the results of other similar experiments. The results of a paired t test only make sense when the pairs are independent – that whatever factor caused a difference (between paired values) to be too high or too low affects only that one pair. Prism cannot test this assumption. You must think about

the experimental design. For example, the errors are not independent if you have six pairs of values, but these were obtained from three animals, with duplicate Measurements in each animal. In this case, some factor may cause the after-before differences from one animal to be high or low. This factor would affect two of the pairs, so they are not independent. The values used for paired and unpaired tests are absolute values.

Unpaired t test results

1. P value and statistical significance:

Two-tailed P value= 0.0115

By conventional criteria, this difference is considered to be statistically significant.

Confidence interval:

The mean of "Touch" minus "Speech with Text"= -0.07047

95% confidence interval of this difference: From -0.12507 to -0.01587

Intermediate values used in calculations:

t = 2.5363

df = 454

Standard error of difference = 0.028

Group	Touch	Speech with Text
Mean	0.27484	0.34531
SD	0.28865	0.30438
SEM	0.01859	0.02076
N	241	215

2. No Distraction vs Touch Mode

P value and statistical significance:

Two-tailed P value= 0.0004

By conventional criteria, this difference is considered to be extremely statistically significant.

Confidence interval:

The mean of "No Distraction" minus "Touch"= 0.10033

95% confidence interval of this difference: From 0.04546 to 0.15520

Intermediate values used in calculations:

$t = 3.5936$ $df = 446$ Standard error of difference = 0.028

Group	No Distraction	Touch
Mean	0.37517	0.27484
SD	0.30142	0.28865
SEM	0.02095	0.01859
N	207	241

3. “No Distraction” vs “Speech with Text”

P value and statistical significance:

Two-tailed P value= 0.3120

By conventional criteria, this difference is considered to be not statistically significant.

Confidence interval:

The mean of “No Distraction” minus “Speech with Text”= 0.02986

95% confidence interval of this difference: From -0.02812 to 0.08785

Intermediate values used in calculations:

$t = 1.0123$ $df = 420$ Standard error of difference = 0.029

Group	No Distraction	Speech with Text
Mean	0.37517	0.34531
SD	0.30142	0.30438
SEM	0.02095	0.02076
N	207	215

One sample t-test

1. “No distraction”:

P value and statistical significance:

Two-tailed P value < 0.0001

By conventional criteria, this difference is considered to be extremely statistically significant.

Confidence interval:

Hypothetical mean= 0.00000 Actual mean= 0.11077

The difference between these two values= 0.11077

The 95% confidence interval of this difference: From 0.06214 to 0.15939

Intermediate values used in calculations:

t = 4.4873 df = 240 Standard error of difference = 0.025

Mean= 0.11077 SD= 0.38321 SEM= 0.02468 N= 241

2. "Touch":

P value and statistical significance:

Two-tailed P value= 0.0446

By conventional criteria, this difference is considered to be statistically significant.

Confidence interval:

Hypothetical mean= 0.00000 Actual mean= 0.06290

The difference between these two values is 0.06290

The 95% confidence interval of this difference: From 0.00152 to 0.12428

Intermediate values used in calculations:

t = 2.0200 df = 214 Standard error of difference = 0.031

Mean= 0.06290 SD= 0.45659 SEM= 0.03114 N= 215

3. "Speech with Text":

P value and statistical significance:

Two-tailed P value < 0.0001

By conventional criteria, this difference is considered to be extremely statistically significant.

Confidence interval:

Hypothetical mean= 0.00000 Actual mean= 0.11077

The difference between these two values is 0.11077

The 95% confidence interval of this difference: From 0.06214 to 0.15939

Intermediate values used in calculations:

t = 4.4873 df = 240 Standard error of difference = 0.025

Mean=0.11077 SD=0.38321 SEM=0.02468 N=241

APPENDIX D

QUESTIONNAIRE FOR THE OUTREACH EVENT

1. What is your
Age: _____
Gender: a. **Male** b. **Female**
2. How familiar are you with driving?
a. **None** b. **Low** c. **Medium** d. **High**
3. Do you know the legal age to get driver's license?

4. Do you use social media? If yes, then what and how often?

5. Do you text? If yes, then how often?

6. Are you familiar with driver distraction? What is it?

7. What can be the possible causes for driver's distraction?

8. Should a drive use cell phone while driving?

9. Should a drive text while driving?

10. Do you think a driver should change the music while driving? _____
11. Which tablet would you prefer to use?
A **B** **C**
D **E** **F**

APPENDIX E
READINGS FOR EXPERIMENT 1

Feature/ Mode	Blind Single Mode Speech				Dual Mode Speech with Text Tip				
	Once	2-3	Failure	Tried Com- mands	Once	2-3	Failure	Tried Com- mands	
GPS	0	3	10	Naviga- tion/Location/Map	13	0	0		
Media	0	1	12	Music/Songs/ Albums	13	0	0		
Phone	3	5	5	Call/Telephone/ Contacts	13	0	0		
Car	9	2	2	Car/ Automobile	13	0	0		
AC	0	0	13	No commands were given by users	11	2	0		Air Condition- er/ Incorrect detection
Settings	13	0	0		12	1	0		Incorrect de- tection
Dial	0	2	11	Call	13	0	0		
Con- tacts	4	5	4	Contact List/Phone book	13	0	0		
Favor- ites	0	2	11	Favorite/Star/ No clue	9	1	3		Favorite
Maps	0	4	9	Navigation/ Lo- cation/Map/ GPS	11	2	0		Map/ Incorrect detection
Wipers	0	3	10	Speed	13	0	0		

Table 5: Experiment 1 readings

APPENDIX F
READINGS FOR EXPERIMENT 2

	Touch					Speech with Text-tip				
V1	2.01	1.03	0.76	0.61	1.1025	1.87	1.13	0.3	0.62	0.98
V2	1.02	0.59	0.78	0.42	0.7025	0.89	0.61	0.87	0.39	0.69
V3	0.33	0.63	0.84	0.74	0.635	1.35	0.54	0.27	0.42	0.645
V4	1.03	0.74	1.07	1.1'4	0.995	0.84	0.54	0.43	0.67	0.62
V5	0.87	1.04	1.81	0.82	1.135	0.94	0.73	1.07	0.86	0.9
V6	0.55	1.17	0.99	0.82	0.8825	0.63	0.78	0.65	0.41	0.6175
V7	0.89	0.955	0.805	0.9	0.8875	1.41	0.98	0.76	0.36	0.8775
V8	0.8	0.835	0.795	0.7	0.7825	1.32	0.92	0.8	0.54	0.895
V9	0.8	0.73	0.92	0.73	0.795	0.55	0.42	0.59	0.84	0.6
V10	0.77	0.76	0.94	1.53	1	0.68	0.81 5	0.76	0.97	0.8062
V11	1.61	1	1.03	0.93 5	1.1437	1.6	0.67	0.78	0.42	0.8675
V12	0.72	0.58	0.795	0.68 5	0.695	0.52	0.38	0.31	0.43	0.41

Table 6: Experiment 2 readings

APPENDIX G
CONSENT FORM

Title of Investigation: Reducing Driver Distraction using Multimodal Interaction

This document is to certify that I, _____, hereby freely agree to participate as a volunteer in a (research study, experiment, program, etc.) as an authorized part of the educational and research program of the Arizona State University under the supervision of Richa Mittal

- The research project has been fully explained to me by Richa, and I understand this explanation, including what I will be asked to do. A copy of the procedures of this investigation and a description of any risks, discomforts and benefits associated with my participation has been provided and discussed in detail with me.
- I have been given an opportunity to ask questions, and all such questions and inquiries have been answered to my satisfaction.
- I understand that I am free to decline to answer any specific items or questions in interviews or questionnaires.
- I understand that all data will remain confidential with regard to my identity.
- I understand that participation in this research project is voluntary and not a requirement or a condition for being the recipient of benefits or services from the Arizona State University or any other organization sponsoring the research project.
- I understand that the approximate length of time required for participation in this research project is (15 minutes maximum).
- I understand that if I have any questions concerning the purposes or the procedures associated with this research project, I may email to rmittal5@asu.edu
- I understand that it will not be necessary to reveal my name in order to obtain additional information about this research project from the principal investigator(s).
- I understand that if I have any questions or concerns about the treatment of human subjects in this study, I may email to rmittal5@asu.edu

- Although this person will ask my name, I understand that all inquiries will be kept in the strictest confidence.

I UNDERSTAND THAT I AM FREE TO WITHDRAW MY CONSENT AND DISCONTINUE MY PARTICIPATION AT ANY TIME.

Date _____

Signature of Subject