Visual and Cognitive Demands of Using Apple's CarPlay, Google's Android Auto and Five Different OEM Infotainment Systems

June 2018

AAA Foundation for Traffic Safety

607 14th Street, NW, Suite 201 | Washington, DC 20005 | 202-638-5944

Title

Visual and Cognitive Demands of Using Apple's CarPlay, Google's Android Auto and Five Different OEM Infotainment Systems

(June 2018)

Authors

David L. Strayer, Joel M. Cooper, Madeleine M. McCarty, Douglas J. Getty, Camille L. Wheatley, Connor J. Motzkus, Kelly L. Mackenzie, Sydney M. Loveless, Jess Esplin, Rachel M. Goethe, and Francesco Biondi

University of Utah

Foreword

The expansion of new infotainment and In-Vehicle Information Systems (IVIS) into vehicles in recent years has afforded drivers new activities and connectivity that can potentially impact safety. It is important to understand how these new technologies impact drivers' workload and performance. Recent work sponsored by the AAA Foundation for Traffic Safety led to the development of new methods for measuring the visual and cognitive demands associated with different in-vehicle systems.

This report expands on earlier efforts from AAA Foundation for Traffic Safety, describing the results of an on-road study looking at the visual and cognitive demand as well as the task completion time for a variety of infotainment tasks and interaction methods. Importantly, the report compares the performance of native OEM infotainment systems in five 2017 model year vehicles with the performance of Apple CarPlay and Android Auto—two popular third-party systems that can be paired with a vehicle's interface. This report and its outcomes should be a useful reference for OEMs, developers of advanced IVIS, public agencies and researchers, as well as the general driving population.

C. Y. David Yang, Ph.D.

Executive Director
AAA Foundation for Traffic Safety

About the Sponsor

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 201
Washington, D.C. 20005
202-638-5944
www.aaafoundation.org

Founded in 1947, the AAA Foundation for Traffic Safety in Washington, D.C., is a not-for-profit, publicly supported charitable research and education organization dedicated to saving lives by preventing traffic crashes and reducing injuries when crashes occur. Funding for this report was provided by voluntary contributions from AAA/CAA and their affiliated motor clubs, individual members, AAA-affiliated insurance companies, and other organizations or sources.

Table of Contents

List of Figures

Terms and Definitions[1]

Center stack – The center stack is located in the center of the dash to the right of the driver. An LCD display is used to present textual and/or graphical information. Center stack systems often include a touch-screen interface to support visual/manual interactions so that drivers can select an option and navigate menus by touch and/or use slider bars to scroll through options displayed on the screen. With some vehicles, the selection of options may be made with manual buttons surrounding the touch screen.

Cohen's d – An effect size estimate derived by a standardized difference between means. A Cohen's d value of 0.2 reflects a small effect size, a value of 0.4 reflects a medium effect size, and a value of 0.8 reflects a large effect size.

Cognitive demand – The cognitive workload associated with the performance of a task. This would include perception, attention, memory and decision-making processes. In this report, we refer to the cognitive demand associated with performing IVIS task types with different modes of interaction when the vehicle is in motion.

Cognitive referent task – The N-back task (see below) served as the cognitive referent task in the current research.

Overall demand – Total visual, auditory, cognitive or physical resources required of the driver to accomplish the primary driving task and interact with an in-vehicle infotainment system in a dual-task setting.

Distraction potential – The potential distraction associated with secondary-task engagement. This potential may not be realized if drivers limit their secondary-task interactions to periods when the vehicle is not in motion.

Driver distraction – The diversion of attention away from activities critical for safe driving toward a competing activity, which may result in insufficient or no attention to activities critical for safe driving (Regan, Hallett and Gordon, 2011).

DRT – The Detection Response Task (DRT) is an International Standards Organization protocol (ISO 17488, 2015) for measuring attentional effects of cognitive load in driving. In this research, a vibrotactile device emitted a small vibration stimulus, similar to a vibrating cell phone, or an LED light stimulus changed color from orange to red. These changes cued the participant to respond as quickly as possible by pressing the microswitch attached to a finger against the steering wheel. DRT reaction time increases and hit rate decreases as the workload of the driver increases.

Dual task – Two tasks performed concurrently, typically the primary driving task plus a secondary task.

Evaluation – A procedure for assessing the effects of an interaction.

In-vehicle information system (IVIS) – The collection of features and functions in vehicles that allow motorists to complete tasks unrelated to driving while operating the vehicle. The IVIS

[1] Some of these terms, definitions and abbreviations were taken directly from ISO (2012), Regan, Hallett and Gordon (2011), and NHTSA (2013).

features we tested involved up to four task types (audio entertainment, calling and dialing, text messaging, and navigation).

Method – High-level approach to an assessment, based on theory and principles, which implies an underlying rationale in the choice of assessment techniques.

Metric **–** Quantitative measure of driver behavior independent of the tool used to measure it.

Linear mixed effects model – We compared the likelihood ratio of the full linear mixed effects model to a partial linear mixed effects model without the effect (e.g., Task, Mode, Task by Mode, Vehicle) to determine if the effect in question accounted for a significant proportion of variance.

NASA TLX – A questionnaire-based metric assessing the subjective workload of the driver. The TLX assesses mental demand, physical demand, temporal demand, performance, effort and frustration.

N-back task – The N-back task presented a prerecorded, randomized set of numbers ranging from zero to nine and presented in sequences of 10. In each sequence, numbers were spoken aloud at a rate of one digit every 2.25 seconds. Participants were instructed to verbally repeat the number that was presented two trials earlier as they concurrently listened for the next number in the sequence. The N-back task places a high level of cognitive demand on the driver without imposing any visual/manual demands.

Performance – The behavior demonstrated by a driver performing the driving task or a related task.

Primary driving tasks – Activities that the driver must undertake while driving, including navigating, path following, maneuvering and avoiding obstacles.

Reference task **–** Type of task used for comparing different tests or test results across vehicles or systems.

Single task baseline – When the driver is performing the primary driving task (i.e., driving) without the addition of workload imposed by IVIS interactions.

Secondary task – A non-driving related additional task.

Secondary task demand – The aggregate of cognitive, visual and manual demands required by a non-driving task.

SuRT task – The variant of the Surrogate Reference Task (SuRT, ISO TS 14198; ISO, 2015) used in this report differed from the ISO standard by requiring participants to use their finger to touch the location of target items (larger circles) presented in a field of distractors (smaller circles) on an iPad Mini tablet computer that was mounted in a similar position in all the vehicles. The SuRT task places a high level of visual/manual demand on the driver because they must look at and touch the display to perform the task. The SuRT task served as a referent for the visual/manual demands associated with performing IVIS interactions.

Task – The process of achieving a specific and measurable goal.

Task interaction time – The time to complete a task. Task interaction time was defined as the time from the moment participants first initiated an action to the time when the last action in the task had terminated and the participant said, “Done”. Based on the project team’s interpretation of the NHTSA visual occlusion test, the maximum total task length should be less than 24 seconds.

Task types – Tasks were categorized into one of four task types: Audio entertainment, calling and dialing, text messaging and navigation, depending on vehicle capabilities. These task types were completed via different modalities equipped in each vehicle (i.e. touch screen, voice recognition) for each interaction.

Visual demand – The visual workload associated with the performance of a task. This would include the structural interference associated with taking the eyes off the forward roadway as well as the central interference in visual processing that arises from cognitive demand. In this report, we refer to the visual demand associated with performing IVIS tasks with different modes of interaction when the vehicle is in motion.

Visual referent task – A variant of the SuRT task (see above) served as the visual referent task in the current research.

Workload – The aggregate of cognitive, visual and manual demands on the driver. A motorist's workload reflects a combination of demands from the primary task of driving and any secondary tasks performed by the driver. The terms "demand" and "workload" are used interchangeably in this report and we develop separate metrics for cognitive workload and visual workload.

Abbreviated Terms

CAMP	Crash Avoidance Metrics Partnership
DRT	Detection Response Task
HASTE	Human Machine Interface and the Safety of Traffic in Europe
IRB	Institutional Review Board
ISO	International Organization for Standardization
IVIS	In-Vehicle Information System
LCD	Liquid-Crystal Display
LED	Light-Emitting Diode
NHTSA	National Highway Traffic Safety Administration
OEM	Original Equipment Manufacturer
OS	Operating System
SAE	Society of Automotive Engineers
SuRT	Surrogate Reference Task
TEORT	Total Eyes-Off-Road Time
USB	Universal Serial Bus

Executive Summary

Many *In-Vehicle Information Systems* (IVIS), also known as infotainment systems, involve complex and multimodal interactions to perform a task. These IVIS interactions may distract motorists from the primary task of driving by diverting the eyes, hands and/or mind from the roadway. Research previously published by AAA Foundation for Traffic Safety provided comprehensive and empirically derived evidence that the workload experienced by drivers systematically varied as a function of the different tasks, modes of interaction and vehicles that were evaluated. This previous assessment suggested that many of these IVIS features were too distracting to be enabled while the vehicle is in motion. However, a growing trend is to provide access to nomadic systems that support various IVIS interactions. For example, both Apple's CarPlay® and Google's Android Auto® are software platforms on the iPhone and Android smartphones, respectively, which allow the driver to pair their phone with the vehicle to perform many of the same tasks offered by the OEMs' systems. It is currently unknown how these hybrid systems perform relative to the native IVIS systems developed by the OEMs.

This report summarizes an on-road study using measures of cognitive demand, visual/manual demand, a subjective workload measure and a measure of the time it took to complete the different tasks using these hybrid systems. The research involved an on-road evaluation of both CarPlay and Android Auto in five different vehicles as participants performed a series of tasks using different modes of interaction. These systems are often marketed as being easier to use than the native systems. Are they? Do the systems vary in different vehicles? How do they compare to each other? How do they compare with the demand of the systems designed by the OEMs?

The results from this research suggested that both CarPlay and Android Auto systems were significantly less demanding than the native OEM infotainment systems for the tasks employed in the study. The strengths and weaknesses of the systems traded off in such a way that the overall demand of CarPlay and Android Auto did not differ. For example, the demand associated with CarPlay was lower with center stack interactions than for auditory/vocal interactions. By contrast, the demand associated with Android Auto was lower with auditory/vocal interactions than for center stack interactions. Similarly, CarPlay had lower overall demand than Android Auto for sending text messages, whereas Android Auto had lower overall demand than CarPlay for destination entry to support navigation. The hybrid systems also varied in demand when they were deployed in different vehicles.

Overall, CarPlay and Android Auto provided more functionality and resulted in lower levels of workload than the native OEM systems. However, both systems had moderately high levels of demand with each often having strengths where the other has weaknesses, providing the opportunity for both to improve the user experience.

Introduction

Driver distraction arises from a combination of three sources (Strayer, Watson, & Drews, 2011). Impairments to driving can be caused by competition for visual information processing, such as when drivers take their eyes off the road to perform a task. Impairments can also come from manual interference, as in cases where drivers take their hands off the steering wheel to perform an operation. Finally, cognitive sources of distraction occur when attention is withdrawn from the processing of information necessary for the safe operation of a motor vehicle. These sources of distraction can operate independently, but they are not mutually exclusive, and therefore different tasks can result in impairments from one or more of these sources. Moreover, few if any tasks are "process pure" (Jacoby, 1991) and instead often place demands on multiple resources (Wickens, 2008).

Many *In-Vehicle Information Systems* (IVIS), also known as infotainment systems, involve complex and multimodal interactions to perform a task. For example, to select a particular music track a driver might push a button on the steering wheel, issue a voice-based command, view options presented on a display located in the center stack and then manually select the desired track by using the touch screen. Complex multimodal IVIS interactions such as this may distract motorists from the primary task of driving by diverting the eyes, hands and/or mind from the roadway (Regan, Hallett, & Gordon, 2011; Regan & Strayer, 2014).

Prior research by Strayer et al. (2017) provided comprehensive and empirically derived evidence that the workload experienced by drivers systematically varied as a function of the different tasks, modes of interaction and vehicles that were evaluated (see also Kidd et al., 2017; Mehler et al., 2015; Zhang et al., 2015; Angell et al., 2006 [Crash Avoidance Metrics Partnership; CAMP] and Engström, Johansson, and Östlund, 2005 [Human Machine interface And the Safety of Traffic in Europe; HASTE]). In fact, the assessment suggested that many of these IVIS features are excessively distracting and should not be enabled while the vehicle is in motion — a finding in line with guidelines developed by the National Highway Traffic Safety Administration (NHTSA, 2013, see p. 24832). However, a growing trend is to provide access to portable systems that support an expansion of various IVIS features and functions. For example, both Apple's CarPlay® and Google's Android Auto® are software platforms on the iPhone and Android smartphones, respectively, that allow the driver to pair their phone with a vehicle to perform many of the tasks offered by the OEMs. It is currently unknown how these integrated, hybrid systems perform relative to the IVIS systems developed by the OEMs.

This report provides a summary of an on-road study where measures of cognitive demand, visual/manual demand, subjective workload and the time it took to complete the different tasks using these hybrid systems were gathered. The native OEM systems were also evaluated using the same tasks. These metrics were evaluated separately and also combined to provide an overall demand score for the different tasks, modes of interaction and vehicles. Following the approaches used by Strayer et al. (2017), the workload metrics were standardized relative to the high demand cognitive referent (i.e., the N-back task had a rating of 1.0) and the high demand visual referent (i.e., the Surrogate Reference Task, SuRT, had a rating of 1.0). Using this integrated workload metric, tasks that had a rating between 0.0 (the demand associated with the single-task baseline) and 1.0 were easier than the high-demand referent and those with ratings greater than 1.0 were harder than the high-demand referent. This procedure provided a rigorous scientific method for directly comparing the workload associated with using CarPlay, Android Auto and native OEM systems.

One component, task duration, is central to the issue of workload assessment. Shutko and Tijerina (2006) suggest that task duration is critical because it represents the cumulative time over which an unexpected event could occur. Based on a model of exposure, they argue that all else being equal, a task that takes twice as long to complete will result in twice the potential risk of an adverse event. Across different studies, task duration is commonly measured independently as a stand-alone performance measure or implicitly as a compound measure (e.g., Reimer et al., 2014, Ito et al., 2001), for example, eyes-off-road time and total task time (see SAE J2944; SAE, 2015). Formally, duration related measures can be defined as measures that co-vary with task duration (Burns et al., 2010). Abstractly, duration related measures are those that involve the accumulation of a measured value over time and can include measurable performance characteristics related to the vehicle control, secondary task performance, driver behavior, etc. A key characteristic of duration-based measures is that they are correlated with total task time and change in value with longer tasks (e.g., longer visual tasks result in greater total eyes-off-road time).

However, there is no clear consensus on what constitutes an acceptable interaction time for a secondary task. The issue is further confounded by research suggesting that secondary tasks are often sensitive to whether testing is completed in a static (i.e., not driving) or dynamic (i.e., driving) environment (Young et al., 2005), the age of participants (McWilliams et al., 2015), and performance characteristics of the primary or secondary tasks (Tsimhoni, Yoo, & Green, 1999). Because of the visual demands associated with driving, visual secondary tasks generally take longer to complete when performed concurrently with driving. Additionally, due to natural aging processes, older adults generally take longer to perform tasks than younger adults.

More recently, NHTSA (2013) has issued a set of voluntary guidelines for visual/manual tasks that suggest that tasks should require no more than 12 seconds of Total Eyes-Off-Road Time (TEORT) to complete (p. 24821). This 12-second guidance is based on the societally acceptable risk associated with tuning an analog in-car radio. Using visual occlusion, a method specified by NHTSA to evaluate visual-manual tasks, motorists can view the driving environment for 12 seconds and vision is occluded for 12 seconds in one-and-one-half-second on/off intervals. The NHTSA guidelines provide a visual occlusion testing procedure — and, using that, the project team derived a task time maximum for use in this project. That derived maximum was 24 seconds (based on the interpretation of the project team that a driver would have 12 seconds of shutter open time + 12 seconds of shutter closed time for a total task time of 24 seconds.)

An important prerequisite for duration-based measures of secondary task performance is the definition of a task. The definition provided by Burns et al. (2010) was used, which suggests that a task can be defined as a sequence of inputs leading to a goal at which the driver will normally persist until the goal is reached. However, the research team differentiates between continuous and discrete tasks that are shaped by different performance goals. Fundamental to secondary discrete tasks is a performance goal with a finite beginning and end state (e.g., changing the audio source, dialing a phone number, calling a contact, entering a destination into a navigation unit, etc.). Conversely, continuous tasks are characterized by performance maintenance over an indefinite period, often with no clear termination state (Schmidt & Lee, 2005). Given the nature of discrete tasks, a failure to account for task duration during assessment provides an incomplete picture of distraction potential.

Experimental Overview

In this report, the research team provides a replication and extension of previous research designed to compare CarPlay and Android Auto when they are used in several different vehicles. These hybrid systems are often marketed as being easier to use than native systems by providing an interface that is more familiar to users and by leveraging the power of remote servers for much of the data-processing requirements. Are these hybrid phone/vehicle interfaces easier to use than native OEM interfaces? Do the hybrid systems vary in different vehicles? How do they compare with each other? How do they compare with the demand of the IVIS systems designed by the OEMs?

The research involved an on-road evaluation of both CarPlay and Android Auto in five different vehicles as participants performed a series of task types using different modes of interaction. The research team also evaluated the demand associated with performing these same tasks using the native OEM systems. Additionally, the single-task, N-back, and SuRT were tested in each vehicle in an experimental order that counterbalanced all experimental conditions across participants. From this design, it was possible to determine the effects of cognitive and visual demand associated with different interface systems (CarPlay, Android Auto and the native OEM systems), task types (calling and dialing, audio entertainment, navigation, and text messaging), and modes of interactions (auditory/vocal vs. center stack).

Method

Participants

Sixty-four participants (32 female) were recruited via flyers and social media posts with approval from the University of Utah Institutional Review Board (IRB). Eligible participants met the following requirements: They were 21-36 years of age (M = 25), were native English speakers, had normal or corrected-to-normal vision, held a valid driver's license and proof of car insurance, and had not been the at-fault driver in an accident within the past two years. To ensure participants held a clean driving record and were eligible to participate in the study, a motor vehicle record report was obtained by the University of Utah's Division of Risk Management. Following University of Utah policy, participants were also required to take and pass a 20-minute online defensive driving course and certification test. Compensation was prorated at $20 per hour.

Twenty-four participants were tested in each configuration of five vehicles crossed with the three different infotainment systems: the native OEM system, CarPlay and Android Auto (i.e., each cell in the 5 X 3 factorial design had 24 participants). A planned missing data design was used (e.g., Graham, Taylor, Olchowski, & Cumsille, 2006; Little & Rhemtulla, 2013) as each participant was tested in an average of two vehicles with CarPlay and two vehicles with Android Auto. Participants were initially naïve to the specific systems and tasks, but were trained until they felt competent and confident performing each type of task.

Stimuli and Apparatus

Vehicles

The vehicles used for the study are listed below with the native infotainment system for each shown in parentheses. Vehicles were selected for inclusion in the study based on whether the vehicle's native system supported both CarPlay and Android Auto, as well as the availability of vehicles for testing. Vehicles were acquired through Enterprise Rental Car or short-term leases from automotive dealerships, or purchased for testing. This sample represents approximately 50% of the platforms that supported both CarPlay and Android Auto in 2017 model year vehicles. The vehicles tested were:

- 2017 Honda Ridgeline RTL-E (HondaLink)
- 2017 Ford Mustang GT (SYNC 3)
- 2018 Chevrolet Silverado LT (MyLink)
- 2018 Kia Optima (UVO)
- 2018 Ram 1500 Laramie (Uconnect)

Equipment and Interaction Systems

Cellular phones on the T-Mobile mobile network were used. Identical LG K7 Android phones were paired via Bluetooth with each vehicle for tasks using the native IVIS systems. Identical iPhone 7 devices were connected via USB to test CarPlay, with software versions iOS 10.3.3 for HondaLink and SYNC 3 and iOS 11.0.3 for UVO, MyLink, and Uconnect. Identical Google Pixel 2 devices connected via USB were utilized to test Android Auto, with software versions OS 7.1.2 for HondaLink and SYNC 3 (all with Android Auto app software version v2.6.573463) and OS

8.0.0 for Uconnect and MyLink (both with Android Auto app software version v2.7.573954) and UVO (with Android Auto app software version v2.8.5754514).

Each vehicle was also equipped with two Garmin Virb XE action cameras, one mounted under the rearview mirror to provide recordings of participants' faces, and another mounted near the passenger seat shoulder to provide a view of the dashboard area for infotainment interaction. Video was recorded at 30 frames per second, at 720p resolution.

An iPad Mini 4 (20.1 cm diagonal LED-backlit Multi-Touch display) was used for the SuRT task and to survey participants on their self-reported measures of workload. Acer R11, Acer Swift and Dell laptop computers were utilized for data collection in the vehicle.

Tasks and Modes of Interaction

During the study, participants interacted with the CarPlay, Android Auto or native OEM system to perform tasks involving audio entertainment, calling and/or dialing, navigation, and text messaging:

- *Audio Entertainment*: Participants changed the music to contents downloaded onto the phone connected via USB. The iPhone 7 and Pixel 2 music contents were identical, but Android Auto uses the Google Play streaming service, whereas Apple CarPlay plays files stored on the phone.
- *Calling and Dialing*: A list of 91 contacts with a mobile and/or work number was created for participant testing. Participants were instructed to call designated contacts and the associated number type was specified when applicable. In vehicles capable of dialing phone numbers, participants were instructed to dial the phone number 801-555-1234 as well as their own phone number.
- *Text Messaging*: Participants were provided with hypothetical scenarios in which they received text messages from various contacts and were instructed to respond via text. Text messages were created by participants using both center stack and auditory vocal modes for each system.
- *Navigation*: Participants started and canceled route guidance to different locations based on hypothetical situations they were given that differed slightly according to the options available in each system.

The modes of interaction with each system are described below. Modes of interaction were selected based on compatibility with the system and individual tasks created based on the vehicle and systems capabilities.

- *Center Stack*: Visual-manual tasks were performed using the center stack interfaces found in the middle of the dash to the right of the driver. Center stack systems generally include a touch screen to integrate visual/manual interactions so that drivers can select options and navigate menus via touch, scroll bars, seek arrows, etc. to complete tasks using options displayed on the screen. Some vehicles provided physical buttons near the touch screen for selection of options.
- *Auditory Vocal*: The voice-based interaction with each third-party system was initiated by the press of a physical voice recognition button on the steering wheel or activated using a vocal command (i.e., "Hey Siri" and "OK Google" commands activated the voice-based assistant for CarPlay and Android Auto, respectively). Microphones

installed in the phone and vehicle process the driver's verbal commands and assist them while performing tasks in the vehicle. Depending on the system, possible voice command options may be presented audibly or displayed on the vehicle's center stack to assist users in achieving their goal.

As noted, the method by which participants interacted with the system depended on the system interface. All vehicles supported voice recognition; however, vehicles differed on the specific visual/manual interaction (e.g., touch screen, manual buttons). Moreover, because they supported different combinations of features and functions, different systems required commands to be given in a specific order and syntax to accomplish tasks in different interaction modes. Task lists were developed in consideration of these differences in order to test the various combinations of features and functions available in each system. Task lists were standardized across systems as much as possible, given the variability in system interactions. The task lists for CarPlay, Android Auto and each native vehicle system are noted in Table 1 and described in detail in Appendix 1. The unique aspects of Android Auto and CarPlay systems as implemented in each vehicle are specified in Appendix 2. (Also, Appendix 3 includes a summary of some of the noteworthy glitches and design concerns that emerged over the course of the testing and evaluation of the systems.)

Table 1. A listing of the tasks and modes of interaction tested in each vehicle. The letters (e.g., A, B, C, etc.) refer to the specific task set instructions described in Appendix 1. Column headers refer to the different tasks by mode of interaction combinations (e.g., AE CS refers to audio entertainment performed using the center stack). Empty cells indicate tasks that were not available for the vehicle/system for that task condition.

Vehicle / System	**Condition**							
	AE CS	**AE AV**	**CD CS**	**CD AV**	**TXT CS**	**TXT AV**	**NAV CS**	**NAV AV**
Third-Party Systems								
Android Auto	C	B	G	G		K	H	I
Apple CarPlay	A	B	F	G	J	K	H	I
Native Systems								
Chevrolet Silverado LT	D	E	G	G	M			
Ford Mustang GT	D	D	G	G	M	L	I	I
Honda Ridgeline RTL-E	D	D	F	G	N		I	I
Kia Optima LX	D	D	G	G				
Ram 1500 Laramie	D	D	G	G		K		

Note: Tasks: AE = Auditory Entertainment; CD = Calling and Dialing; TXT = Text Messaging; NAV = Destination Entry using Navigation. Mode of Interaction: CS = Center Stack; AV = Auditory Vocal.

Detection Response Task (DRT)

Participants responded to both a vibrotactile stimulus and a remote visual stimulus (cf. ISO 17488; ISO, 2015). A vibrotactile device was positioned under the participant's left collarbone and, following ISO guidelines, the vibrotactile device emitted a small vibration stimulus intermittently, similar to a vibrating cell phone. A remote LED light was also placed along a strip of fabric fastener on the dashboard, such that the participants only saw the reflection of the light, directly in their line of sight. The remote light stimulus consisted of a change in color from orange to red, a variant from the ISO standard, developed and evaluated by Castro, Cooper, & Strayer (2016; see also Cooper, Castro, & Strayer, 2016).

A microswitch was attached to either the index or middle finger of the left hand and pressed against the steering wheel when participants felt a vibration or saw the light change colors. The occurrences of these stimuli cued the participant to respond as quickly as possible by pressing the microswitch against the steering wheel. The tactor and light were equally probable and programmed to occur every three to five seconds (with a rectangular distribution of interstimulus intervals within that range) and lasted for one second or until the participant pressed the microswitch. Each press of the switch was counted and recorded but only the first response was used to determine response time to the stimulus. Millisecond resolution response time to the vibrotactile onset or LED light was recorded using a dedicated microprocessor that passed results over USB connection to the host computer for storage and later analysis.

This variant of the standard DRT was used to maximize sensitivity to both cognitive and visual attention. Reaction time to the vibrotactile stimulus was used to assess cognitive load, while hit rate to the forward LED was used as a measure of competing visual demand.

Procedure

Participants completed specific steps involving interactions with CarPlay or Android Auto, or native systems to complete a task (i.e., using the touch screen to tune the radio to a particular station, using voice recognition to find a particular navigation destination, etc.) while driving the vehicle. Prior to task training, each participant completed the respective voice trainings for "Hey Siri" or "OK Google" on the iPhone 7 or the Pixel 2. Tasks were categorized into one of four task types: audio entertainment, calling and dialing, text messaging, and navigation. These task types were completed using different interfaces equipped in each vehicle (i.e., center stack or voice recognition) for each interaction (see Table 1). The order of interactions was counterbalanced across participants.

Driving Route

A suburban residential street with a 25 mph speed limit was used for the on-road driving study. The route consisted of a straight road with four stop signs and two speed bumps. Participants were required to follow all traffic laws and adhere to the speed limit at all times. The driving route was approximately two miles long one-way with an average drive time of six minutes. A researcher was present in the passenger seat of each vehicle for safety monitoring and data collection.

Training

Prior to the start of the study, participants were provided the time to become accustomed to the vehicle, the route and the DRT prior to data collection. The initial familiarization period included:

- *Practice Route:* Participants were instructed to drive the assigned route, while researchers pointed out all obvious and identifiable road hazards.
- *DRT Training:* Once the participants felt capable driving the vehicle, they were trained to respond to the DRT (vibrotactile and LED light). Researchers monitored participants as they practiced responding to 10 stimuli presented between three to five seconds apart to ensure participants produced response times of less than 500 milliseconds, indicating a competence and understanding of the task.
- *Voice Pairing (Apple CarPlay and Android Auto Only):* To improve voice activation and accuracy, participants completed the voice training specific to each system via "Hey Siri" and "OK Google" voice commands. The native systems tested did not allow for this unique voice pairing process.

Participants were trained to interact with and complete the tasks using the assigned mode of interaction before each condition began. Participants were required to complete three task trials without error prior to starting the driving task for each of the system interactions. Once participants demonstrated competence in their ability to interact with the system, the experiment began.

Experimental Blocks

During the experimental blocks, participants were instructed to complete a set of tasks administered by the researcher using an assigned mode of interaction with the infotainment system. Driving the vehicle was considered the primary task, interacting with the infotainment system was considered the secondary task and responding to the DRT was considered the tertiary task.

At each end of the route, participants were asked to pull over on the side of the road, indicating the conclusion of the experimental block. The following experimental block, which included a new assigned task and mode of interaction, began in the opposite direction of the same route and concluded in the same manner. This was repeated until all conditions were completed, resulting in alternating travel directions for each experimental block.

Tasks were only administered in safe and normal driving conditions. Disruptions to the natural driving environment resulted in the researcher instructing the participant to terminate the current task and only administering a new task when it was safe to do so. Tasks were not administered as participants approached hazards such as intersections and construction zones. Other hazards that may have impeded safe driving behavior, such as pedestrians or cyclists causing the driver to leave their designated lane or slow more than five miles below the speed limit, resulted in longer "off task" periods or required task termination. Behaviors of other vehicles and pedestrians were largely considered normal and within the scope of the natural driving environment.

For the tasks, participants were provided with verbal hypothetical situations or commands as cues from the researcher (e.g., "You want to hear the Essential Johnny Cash album"). Participants were instructed to wait to start each task until the researcher said, "Go." After the completion of each task, participants were trained to say, "Done." Tasks were delivered with a five second interval between the participants' announcement of completion and the researcher's administration of the next task. The researcher denoted each task's start and end time by pressing designated keys on the data collection computer, thus indicating the timing of on-task performance on the driving route. DRT trials were considered valid for statistical analysis if they occurred between these start and end times. Participants were encouraged to complete tasks as efficiently as possible; however, they were given as much time as needed to complete each task, unless the end of the route was reached. In these cases, tasks were terminated prematurely and later omitted from analysis. The total number of tasks administered and completed in each two-mile run thus varied depending on task duration.

Participants also performed three control tasks while driving one length of the designated route per task. The control tasks were:

- *Single-task baseline*: Participants performed a single-task baseline drive on the designated route, without interacting with the infotainment system. During the single-task baseline, participants interacted solely with the DRT stimuli, responding to both the tactor and light change as fast as possible, and were asked to remain silent as to minimize distraction.
- *Auditory N-back task*: The auditory N-back task presented a prerecorded, randomized set of numbers ranging from zero to nine in sequences of 10 (e.g., Mehler et al., 2011). In each sequence, numbers were spoken aloud at a rate of one digit every 2.25 seconds. Participants were instructed to verbally repeat the number that was presented two trials earlier as they concurrently listened for the next number in the sequence. Participants were told to respond as accurately as possible to the N-back stimuli while researchers monitored performance in real time. During the N-back task, participants also responded to the DRT stimuli.
- *SuRT task*: The Surrogate Reference Task (SuRT) was presented on an iPad Mini 4 with circles printed in black on a white background. A target was presented on the display amid 21-27 distractors. The target was an open circle measuring 1.5 cm in diameter and the distractors were open circles measuring 1.2 cm in diameter. For each trial, participants were instructed to touch the location of the target.[2] Immediately thereafter, a new display was presented with a different configuration of targets and distractors. The location of targets and distractors was randomized across the trials in the SuRT task. Participants were instructed to continuously perform the SuRT task while giving the driving task highest priority as researchers monitored performance in real time. Researchers instructed participants to pause the SuRT task at intersections and in the event of potential hazards on the roadway. During the SuRT task, participants also responded to the DRT stimuli.

After the completion of each condition, participants completed the NASA-TLX (Hart & Staveland, 1988) to assess the subjective workload of the system and were given the opportunity to make comments about the task using a form presented on the iPad Mini 4.

Dependent Measures

DRT data were processed following procedures outlined in ISO 17488 (ISO, 2015). All response times faster than 100 milliseconds or slower than 2500 ms were eliminated from our overall calculation for reaction time. Nonresponses or responses that were made after 2.5 seconds from the stimulus onset were coded as misses. System interactions (i.e., tasks) were coded by the researcher by pressing designated keys on the DRT host computer, allowing the identification of "on-task" and "off-task" segments of driving. Incomplete, interrupted or otherwise invalid tasks

[2] The variant of the SuRT task used in the current research matched as closely as possible the visual display characteristics described in ISO/TS 14198 (ISO, 2012); however, participants responded to the target by pressing the touch-screen location rather than using a directional keypad. This task places visual/manual demands on drivers that are more similar in nature to interactions using the center stack LCD touch screen.

were flagged and excluded from the analysis. The DRT-related dependent measures used in the study are described below:

- DRT – Reaction Time: Defined as the sum of all valid reaction times to the DRT task divided by the number of valid reaction times.
- DRT – Hit Rate: Defined as the number of valid responses divided by the total number of valid stimuli presented during each condition.

Upon the completion of each condition, participants were asked to complete a brief survey that identified eight questions related to the task. The first six of these questions were from the NASA TLX; the final two assessed the intuitiveness and complexity of the IVIS interactions. Reponses to these subjective measures were made on a 21-point scale for each question:

- Mental – How mentally demanding was the task?
- Physical – How physically demanding was the task?
- Temporal – How hurried or rushed was the pace of the task?
- Performance – How successful were you in accomplishing what you were asked to do?
- Effort – How hard did you have to work to accomplish your level of performance?
- Frustration – How insecure, discouraged, irritated, stressed and annoyed were you?
- Intuitiveness – How intuitive, usable and easy was it to use the system?
- Complexity – How complex, difficult and confusing was it to use the system?

Task interaction time was derived from the time stamp on the DRT data file and defined as the time participants first initiated an action to the time when the final action for a task was completed and the participant said, “Done”. Tasks with irregular occurrences and errors in administration or performance that may have affected task interaction time were marked as abnormal during data collection and were not included in subsequent analyses.

Data Analysis and Modeling

Following the procedures described by Strayer et al. (2017), the DRT data were used to provide empirical estimates of the cognitive and visual demand for the different conditions. To estimate cognitive demand, the average RT to the vibrotactile DRT for each participant was first computed for the single-task baseline condition and for the N-back task (the referent conditions). For all other conditions in the study, Equation 1 was used to standardize the vibrotactile DRT data.

$$Equation\ 1{:}\ Cognitive\ Demand = \frac{IVIS\ Task - Single\ Task}{Nback\ Task - Single\ Task}$$

Using Eq. 1, the single-task baseline receives a rating of 0.0 and the N-back task receives a score of 1.0. It follows that IVIS tasks tested in the vehicle were similarly scaled such that values below 1.0 would represent a cognitive demand lower than the N-back task and values greater than 1.0 would denote conditions with a higher cognitive demand than the N-back task. Cognitive demand is thus a continuous measure ranging from 0 to ∞, with higher values indicating higher levels of cognitive demand.

To estimate visual demand, the average hit rate to the remote DRT for each participant was computed for the single-task baseline condition and for the SuRT task. Equation 2 was then used to standardize the data collected from the remote DRT for all other task conditions.

$$Equation\ 2:\ Visual\ Demand = \frac{Single\ Task - IVIS\ Task}{Single\ Task - SuRT\ Task}$$

Following Eq. 2, the single-task baseline receives a rating of 0.0 and the SuRT task receives a score of 1.0. IVIS tasks tested in the vehicle were similarly scaled such that values below 1.0 would represent visual demand lower than the SuRT task and values greater than 1.0 would denote conditions with visual demand higher than the SuRT task. As with cognitive demand, the visual demand is a continuous measure ranging from 0 to ∞, with higher values indicating higher levels of demand.

To estimate subjective demand, the average of the six NASA TLX subscales for each participant were computed for the single-task baseline condition and for the N-back and SuRT tasks. Equation 3 was used to standardize the subjective estimates.

$$Equation\ 3:\ Subjective\ Demand = \frac{IVIS\ Task -\ Single\ Task}{(\frac{Nback\ Task + SuRT\ Task}{2}) - Single\ Task}$$

Using Eq. 3, the single-task baseline would receive a rating of 0.0 and the average of the N-back and SuRT tasks receives a score of 1.0. IVIS tasks tested in the vehicle were similarly scaled such that values below 1.0 would represent a subjective demand lower than the average of the N-back and SuRT tasks and values greater than 1.0 would denote conditions with subjective demand higher than the average of the N-back and SuRT tasks. As with cognitive and visual demand, the subjective demand is a continuous measure ranging from 0 to ∞, with higher values indicating higher levels of subjective demand.

Equation 4 was used to standardize the IVIS interaction time data using the 24-second interaction time referent.

$$Equation\ 4:\ Interaction\ Time = \frac{IVIS\ Task}{24\ seconds}$$

Using Eq. 4, a task interaction time of 24 seconds receives a score of 1.0. IVIS interactions tested in the vehicle were scaled such that values below 1.0 would represent a task interaction time lower than 24 seconds and values greater than 1.0 would denote conditions with a task interaction time greater than 24 seconds. The time-on-task metric is a continuous measure ranging from 0 to ∞, with higher values indicating longer task interaction time.[3]

An overall workload rating was determined by combining the cognitive, visual and subjective demand with the interaction time rating using Eq. 5. Following from Eq. 5, overall demand is a continuous measure ranging from 0 to ∞, with higher values indicating higher levels of workload.

$$Equation\ 5:\ Overall\ Demand = \frac{(Cognitive\ + Visual\ + Subjective)}{3} * Interaction\ Time$$

[3] As described in earlier sections, the 24-second task interaction referent is derived from NHTSA (2013). The general principle is that these multimodal IVIS interactions should be able to be performed in 24 seconds or less when paired with the task of operating a moving motor vehicle.

Applications of these formulas provide stable workload ratings with useful performance criteria that are grounded in industry standard tasks. On occasion, however, the approach can result in extreme values when either the numerator is unusually small or the task time unusually long. To mitigate the potential for such scores to skew the overall rating, scores greater than 3.5 standard deviations from the mean (<1% of the data) were excluded from analysis.

Experimental Design

The experiment was a 5 (Vehicle) x 3 (System: OEM native system; CarPlay; Android Auto) x 4 (Task Type: audio entertainment; calling and dialing; text messaging; navigation) x 2 (Mode of Interaction: auditory/vocal; center stack) factorial design with 24 participants evaluated in each of the Vehicle x Interaction cells of the factorial. However, not all systems and vehicles offered the full factorial design (see Table 1). Moreover, participants were tested in a varying number of systems. Consequently, an unbalanced design was used where some cells in the factorial were missing because not every participant drove every vehicle. It was necessary to use this approach because it was not practical or feasible for all participants to drive all cars — especially as different vehicles were available at different points in time during the study. Linear mixed models are specifically designed to accommodate this form of missing data (i.e., they are appropriate for unbalanced designs with different numbers of observations for different participants). The number of systems tested by the different participants was used in all linear mixed effects models presented below in order to control for any impact of this latter factor.

Results

A bootstrapping procedure was used to estimate the 95% confidence intervals (CI) around each point estimate in the analyses reported below. The bootstrapping procedure, needed because the standardized scores are ratios derived from other measures, used random sampling with replacement to provide a nonparametric estimate of the sampling distribution. The bootstrapping procedure involved generating 10,000 bootstrapping samples, each of which were created by sampling with replacement N samples from the original "real" data. From each of the bootstrap samples, the mean was computed and the distribution of these means across the 10,000 samples was used to provide an estimate of the standard error around the observed point estimate.[4]

The greater the spread of the CI, the greater the variability associated with the point estimate. The obtained 95% CI also provides a visual depiction of the statistical relationship between the point estimate and the single-task baseline and/or the high demand referents for cognitive, visual, subjective and interaction time. For example, if the high demand referent does not fall within the 95% CI, then the point estimate significantly differs from that referent. Similarly, if the 95% CI of two conditions do not overlap, then the two conditions differ significantly.

The standardized scores for the high demand cognitive or visual referent tasks (N-back and SuRT, respectively) can also be translated into effect size estimates (i.e., Cohen's d). For cognitive demand, a standardized score of 1.0 reflects a Cohen's d of 1.423. For visual demand, a standardized score of 1.0 reflects a Cohen's d of 1.519. Thus, the high demand estimates for cognitive and visual referent tasks reflect *very large* effect sizes. Note that a standardized score of 2 would reflect a doubling of the effect size estimates, a standardized score of 3 would reflect a tripling of the effect size estimates, and so on. Note also that the effect size estimates for the high cognitive and visual demand are virtually equivalent (differing by less than 0.1 Cohen's d units).

Linear mixed effects analyses were performed using R 3.3.1 (R Core Team, 2016), lme4 (Bates, Maechler, Bolker, & Walker, 2015), and multcomp (Hothorn, Bretz, & Westfall, 2008). In the analyses reported below, Task Type, Mode of Interaction, Task Type x Mode of Interaction, and Vehicle were entered independently. The number of vehicles driven by each participant was entered as a fixed effect while Participant, Vehicle, Mode of Interaction and Task Type were entered as random effects. In each case, p-values were obtained by likelihood ratio tests comparing the full linear mixed effects model with a partial linear mixed effects model without the effect in question. This linear mixed modeling analysis has the advantage of analyzing all available data while adjusting fixed effect, random effect and likelihood ratio test estimates for missing data.

The full linear mixed effects model can be expressed as:

$$DV_{ijkl} = B_{0jkl}\text{*intercept} + B_{1jkl}\text{*system} + B_{2jkl}\text{*vehicles run} + e_{0jkl} + m_{00kl} + u_{000l} + r_{ijkl},$$

where *i* indexes task, *j* indexes mode of interaction, *k* indexes vehicles and *l* indexes participants.

[4] Prior to bootstrapping all scores were baseline corrected, minimizing the potential for violations of homogeneity of variance in resampling procedures (e.g., Davidson, Hinkley, & Young, 2003). The baseline correction eliminated any effects of participant in the analyses reported below.

Empirical Data and Inferential Statistics

The empirical data are presented in Figures 1-35. The figures illustrate the major trends from the factorial analysis for the dependent measures of cognitive demand, visual demand, subjective demand, and task interaction time and the integrated overall demand score (based on Eq. 5). A description of the major trends obtained in the linear mixed effects analysis follows the presentation of each dependent measure.

Figures 1-5 present empirical data related to the type of system used to complete the tasks. Each system/interface type supported a similar, though not identical, set of tasks (c.f. Appendix 1). Scores reflect the average demand associated with performing the aggregate task set, collapsed across vehicle, mode of interaction and task.

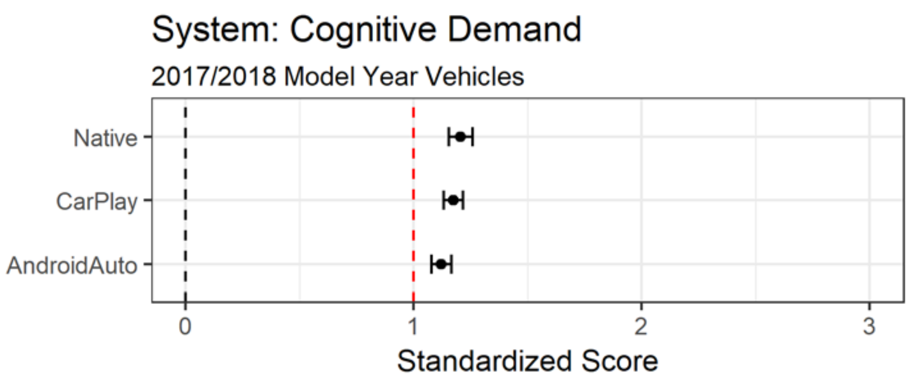

Figure 1. Cognitive demand associated with the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without System indicated that the effect of System was not significant ($X^2(2) = 4.23$, $p > .05$).

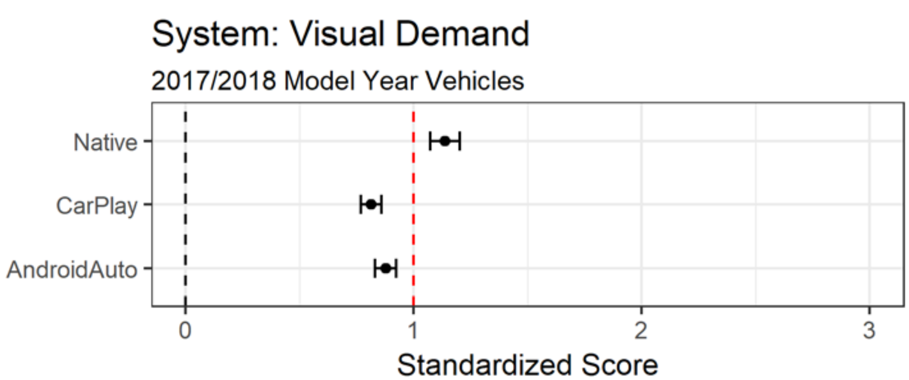

Figure 2. Visual demand associated with the native, CarPlay and Android Auto systems (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without System indicated that the effect of System was significant ($X^2(2) = 9.99$, $p < .01$).

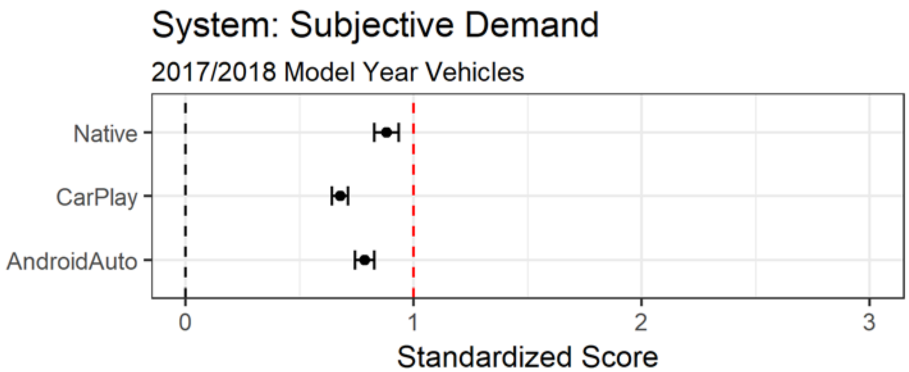

Figure 3. Subjective demand associated with the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without System indicated that the effect of System was significant ($X^2(2) = 6.94$, $p < .05$).

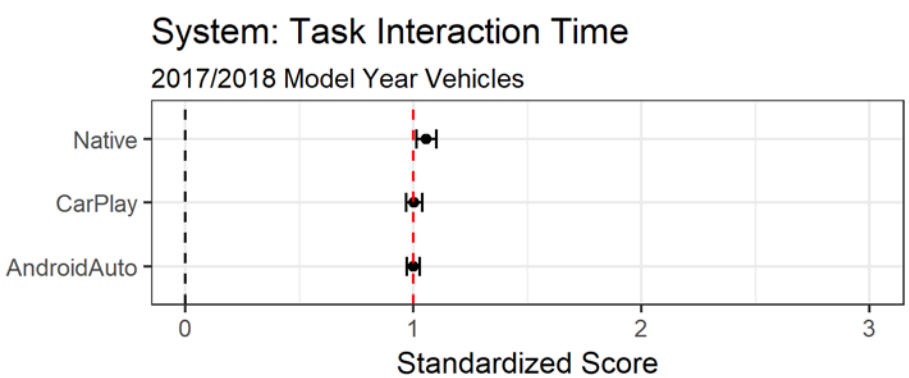

Figure 4. Task interaction time associated with the native, CarPlay and Android Auto system (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without System indicated that the effect of System was not significant ($X^2(2) = 3.62$, $p > .05$).

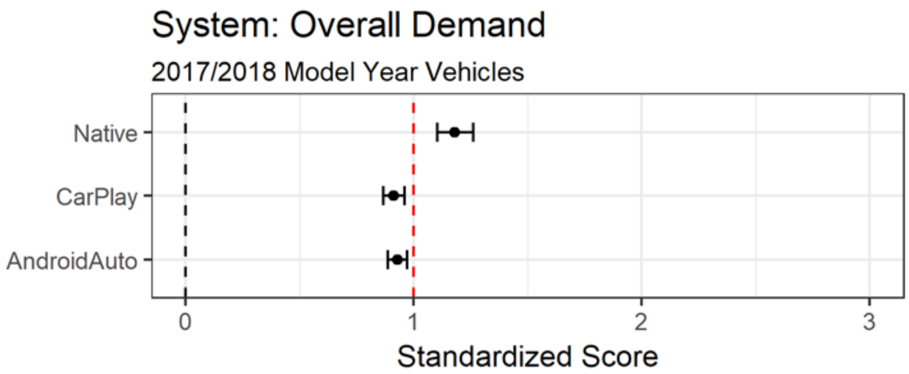

Figure 5. Overall demand associated with the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without System indicated that the effect of System was significant ($X^2(2) = 6.04$, $p < .05$).

Several interesting profiles are noteworthy in Figures 1 through 5. First, cognitive demand across each of the systems was relatively constant and was higher than the N-back reference task. Conversely, both the CarPlay and Android Auto platforms resulted in less visual demand than native systems, and both hybrid systems were significantly below the SuRT reference task. The overall demand of CarPlay and Android Auto systems did not significantly differ from one another; however, both resulted in significantly lower levels of workload than the native systems and both significantly below the red line (referent tasks).

Figures 6 through 10 present the results of the interaction between System and Mode of Interaction for each of the component demand measures. These figures address questions related to how demands for each system differed by interaction mode.

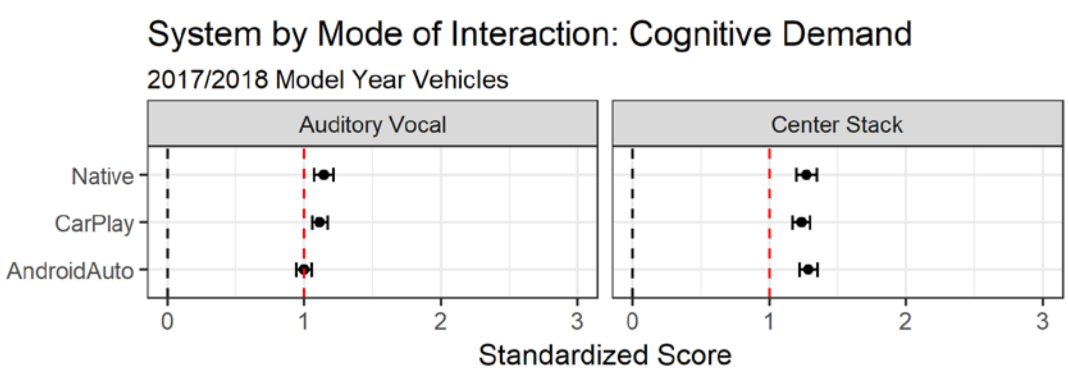

Figure 6. Cognitive demand associated with auditory/vocal and center stack interactions with the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Mode of Interaction indicated that this interaction was significant ($X^2(2) = 15.97$, $p < .01$).

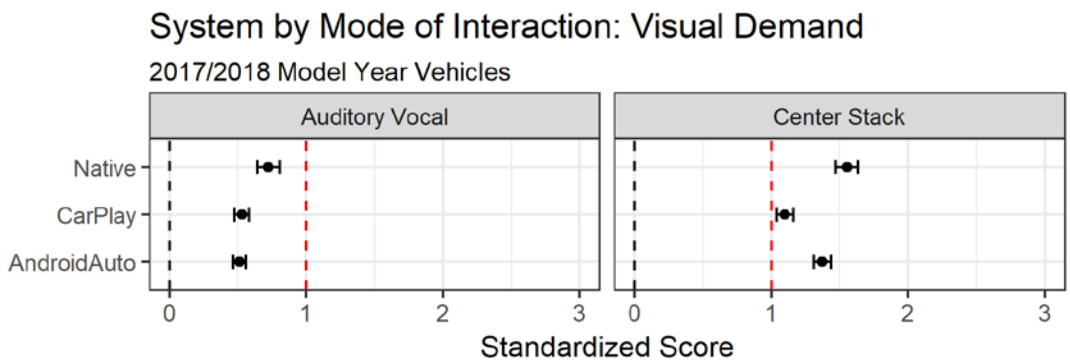

Figure 7. Visual demand associated with auditory/vocal and center stack interactions with the native, CarPlay and Android Auto systems (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Mode of interaction indicated that this interaction was significant ($X^2(2) = 35.86$, $p < .01$).

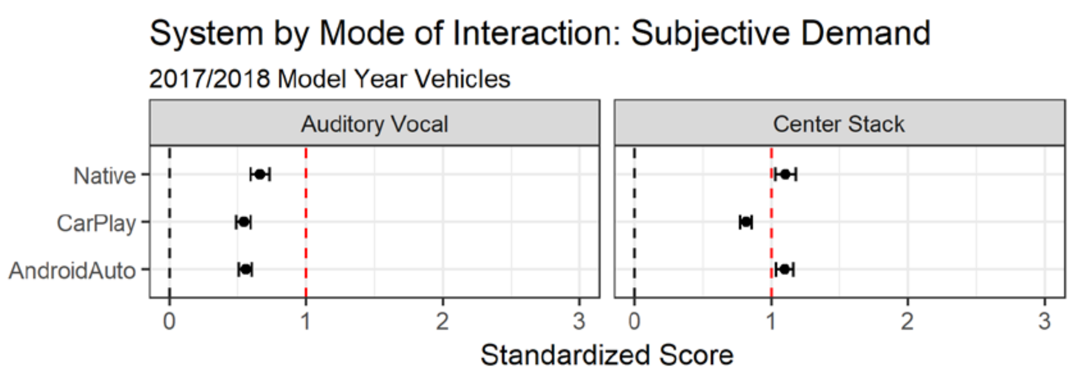

Figure 8. Subjective demand associated with auditory/vocal and center stack interactions with the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Mode of Interaction indicated that this interaction was significant ($X^2(2) = 39.44$, $p < .01$).

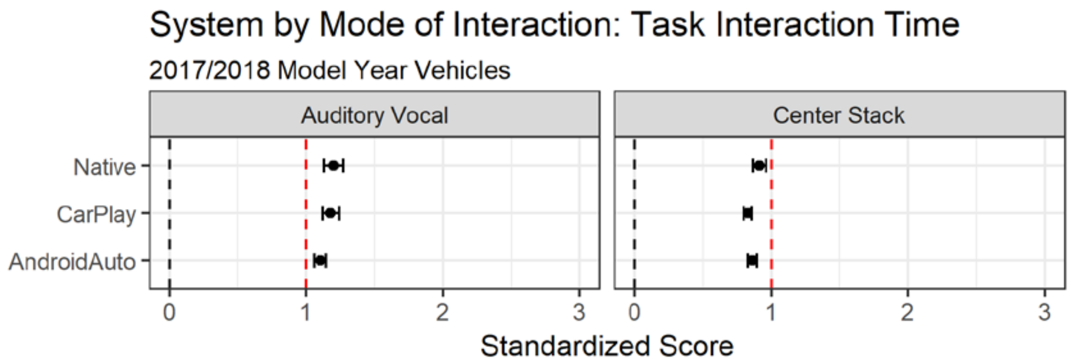

Figure 9. Task interaction time associated with auditory/vocal and center stack interactions with the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Mode of Interaction indicated that this interaction was significant ($X^2(2) = 26.07$, $p < .01$).

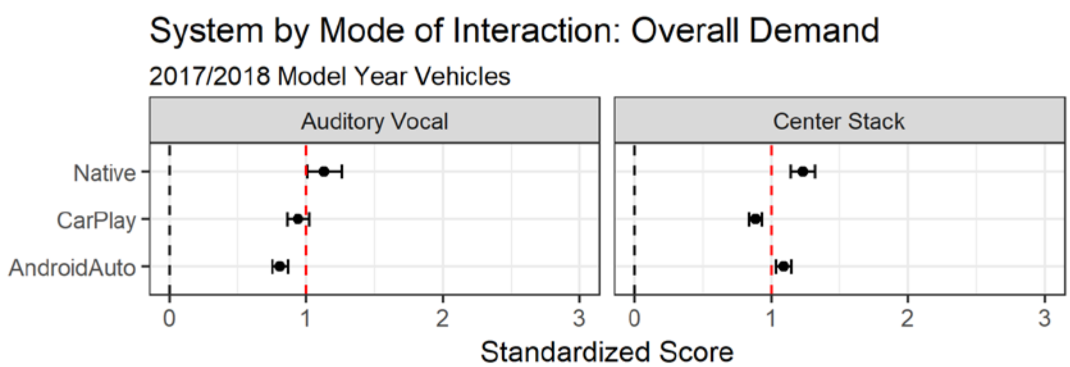

Figure 10. Overall demand associated with auditory/vocal and center stack interactions with the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Mode of Interaction indicated that this interaction was significant ($X^2(2) = 44.33$, $p < .01$).

Several interesting profiles are noteworthy in Figures 6 through 10. Cognitive demand was lower for the auditory/vocal interface than for the center stack interface. This trend was apparent with the CarPlay system and even more pronounced with Android Auto. Visual demand associated with the hybrid systems was lower for both the auditory vocal and center stack interactions when compared with the native systems. Overall demand, for both modes of interaction, was highest with the native OEM systems, followed by CarPlay, and Android Auto. Interestingly, overall demand with CarPlay was lower for center stack interactions than auditory/vocal interactions. By contrast, overall demand for Android Auto was lower for auditory/vocal interactions than for center stack interactions.

Figures 11 through 16 present the workload of each system broken down by task. These data provide insight into possible demand differences of the different tasks for each of the three types of systems.

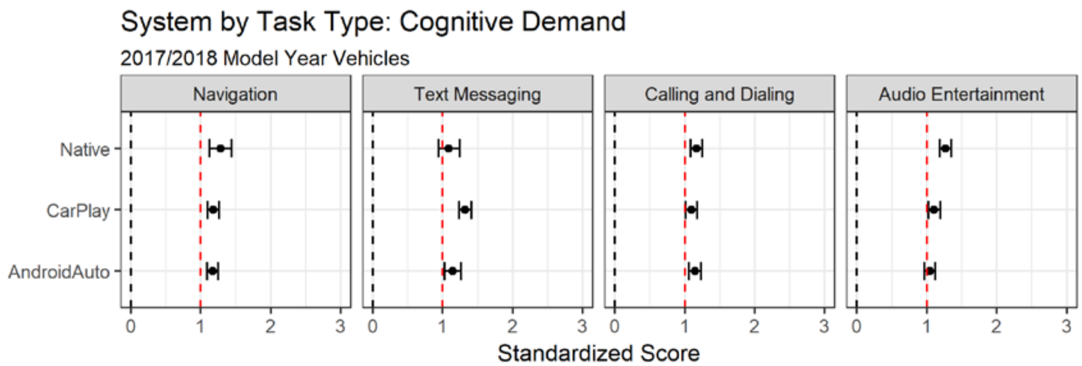

Figure 11. Cognitive demand associated with navigation, text messaging, calling and dialing, and audio entertainment with the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Task interaction indicated that this interaction was significant ($X^2(6) = 24.28$, $p < .01$).

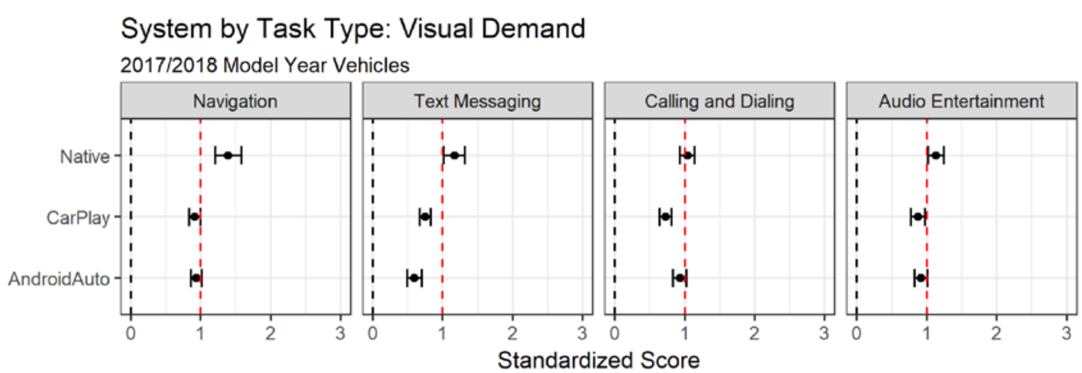

Figure 12. Visual demand associated with navigation, text messaging, calling and dialing, and audio entertainment with the native, CarPlay and Android Auto systems (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Task interaction indicated that this interaction was significant ($X^2(6) = 14.24$, $p < .05$).

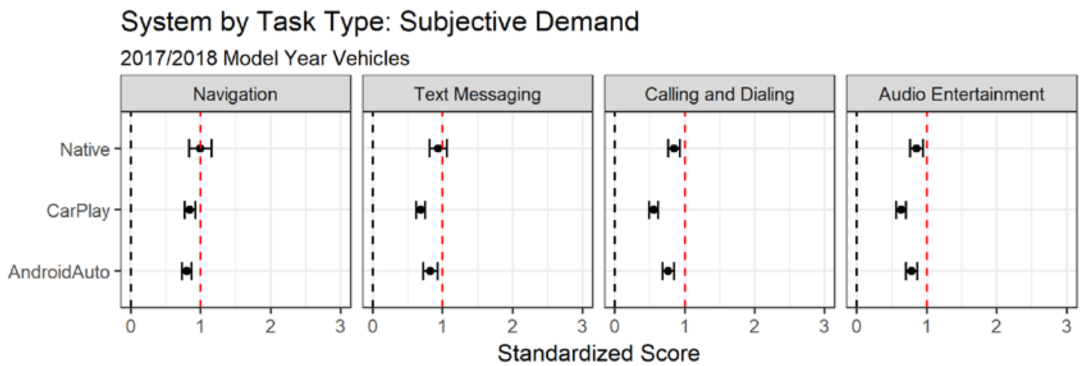

Figure 13. Subjective demand associated with navigation, text messaging, calling and dialing, and audio entertainment with the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Interface by Task interaction indicated that this interaction was significant ($X^2(6) = 31.13$, $p < .01$).

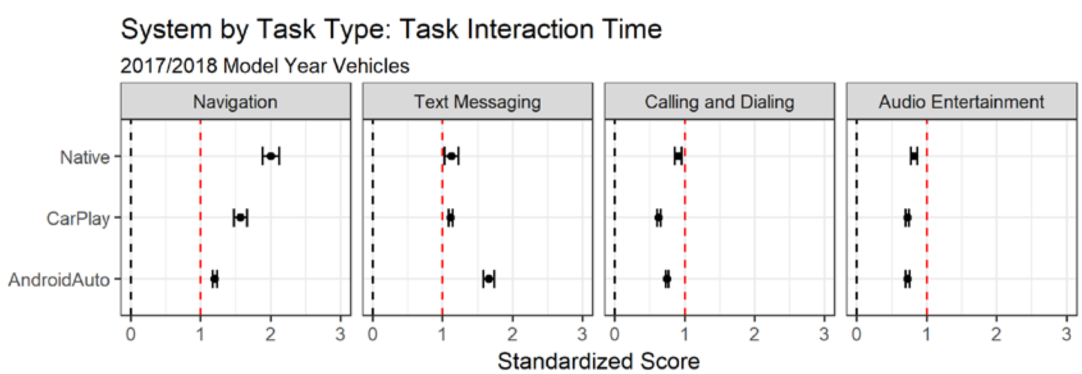

Figure 14. Task interaction time associated with navigation, text messaging, calling and dialing, and audio entertainment with the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Task interaction indicated that this interaction was significant ($X^2(6) = 415.12$, $p < .01$).

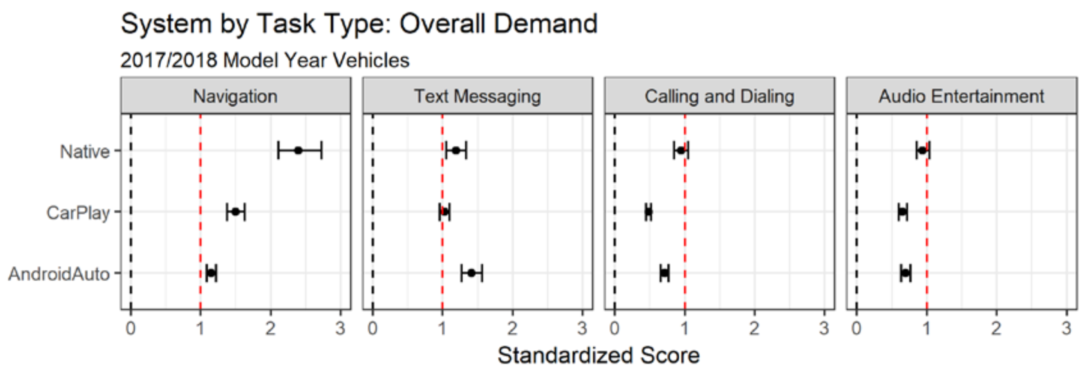

Figure 15. Overall demand associated with navigation, text messaging, calling and dialing, and audio entertainment with the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Task interaction indicated that this interaction was significant ($X^2(6) = 168.98$, $p < .01$).

Analyses of system by task suggest that demand profiles were driven by the type of task being performed. Notably, the time demands required by each task were quite variable across the three interface types. For example, the navigation task took 15 seconds longer to complete for the native systems (M = 48 s) compared to the hybrid systems (M = 33 s). Moreover, the overall demand across task clearly illustrates performance trade-offs. For example, overall demand when sending text messages was lower with CarPlay than it was for Android Auto, but Android Auto had lower overall demand than CarPlay for navigation entry. In most cases, the native OEM systems were associated with higher overall demand than CarPlay and Android Auto. (The exception was text messaging where Android Auto was nominally more demanding than the native OEM system.)

Figures 16 through 20 present the demands of each system by vehicle type. These figures provide insight into the relative demand consistency of Android Auto and CarPlay across the five vehicles.

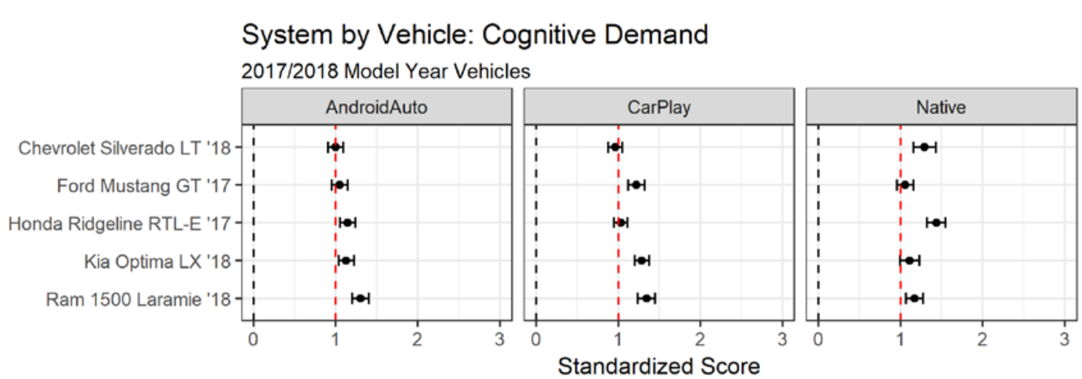

Figure 16. Cognitive demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Vehicle interaction indicated that this interaction was significant ($X^2(8) = 42.67$, $p < .01$).

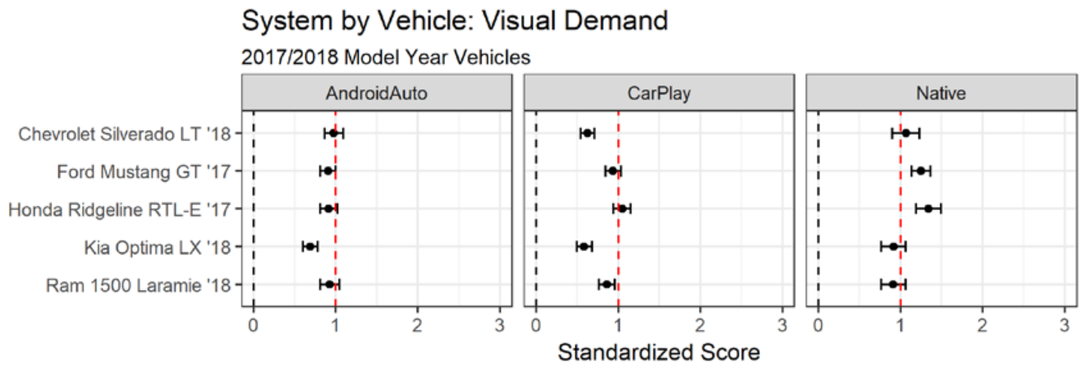

Figure 17. Visual demand associated with five different vehicles using the native, CarPlay and Android Auto systems (from Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Vehicle interaction indicated that this interaction was significant ($X^2(8) = 40.65$, $p < .01$).

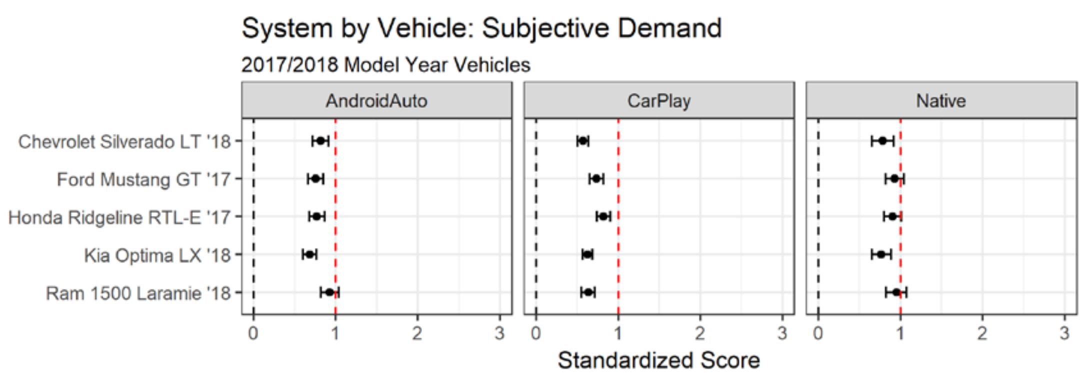

Figure 18. Subjective demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Vehicle interaction indicated that this interaction was significant ($X^2(8) = 35.8$, $p < .01$).

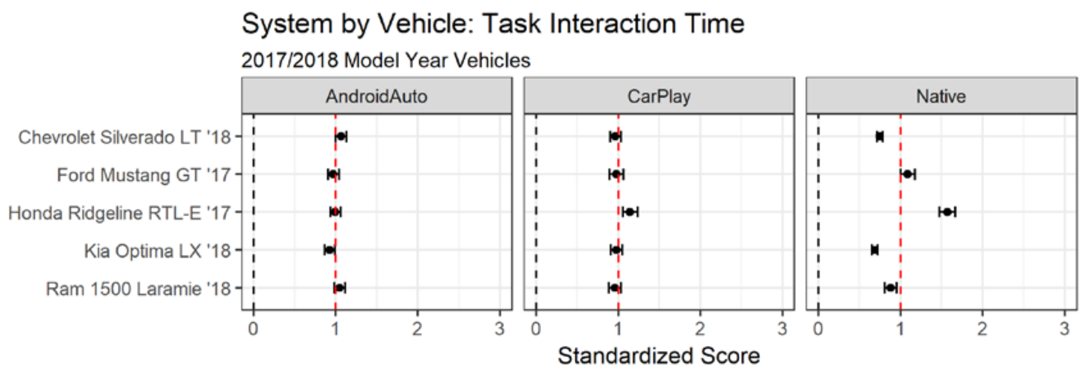

Figure 19. Task interaction time associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Vehicle interaction indicated that this interaction was significant ($X^2(8) = 133.21$, $p < .01$).

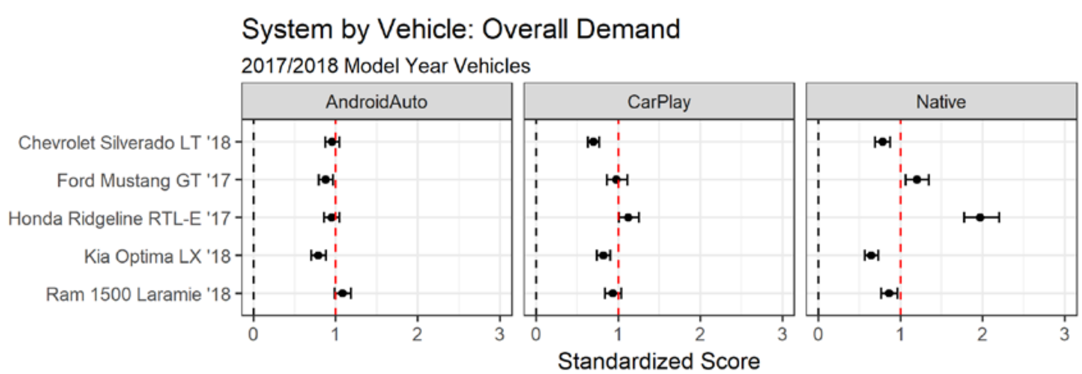

Figure 20. Overall demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the System by Vehicle interaction indicated that this interaction was significant ($X^2(8) = 89.38$, $p < .01$).

A general trend seen in Figures 16 through 20 is that the various components of demand for Android Auto and CarPlay were relatively consistent across vehicles. The overall demand scores suggested that Android Auto was more consistent across vehicles than CarPlay, which was more consistent than the native OEM systems.

Figures 21 through 35 present a number of three-way interactions with System, Vehicle, Mode of Interaction and Task. These figures provide granular results that can be used to identify areas of exceptional performance. Notably, the overall demand associated with using Android Auto and CarPlay differ by vehicle and mode of interaction (see Figure 25). The overall demand for the native systems is more variable, particularly for auditory/vocal interactions. The variability across vehicles for center stack interactions is lower than for auditory/vocal interactions with CarPlay. By contrast, for Android Auto, the variability is greater across vehicles for center stack interactions than for auditory/vocal interactions.

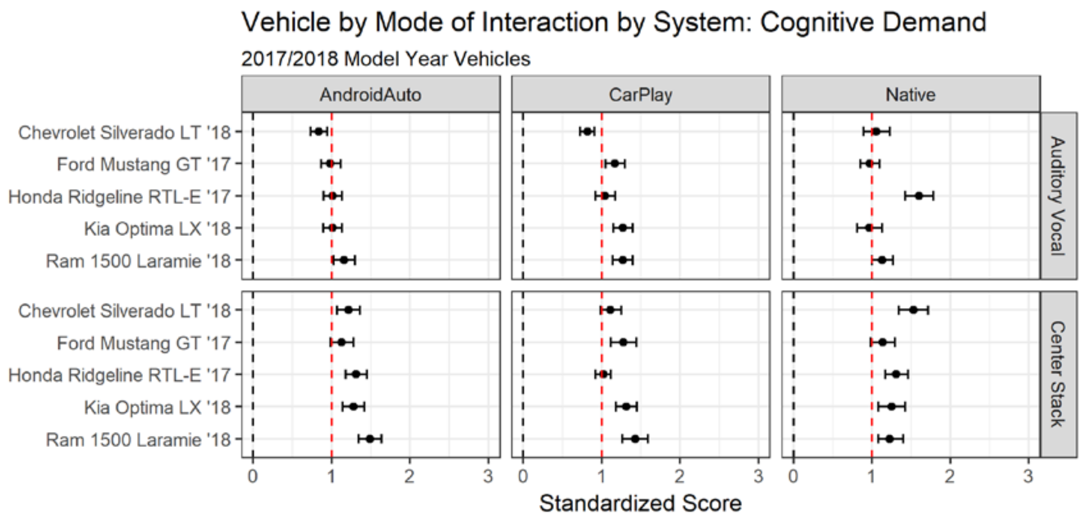

Figure 21. Cognitive demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(22) = 108.39$, $p < .01$).

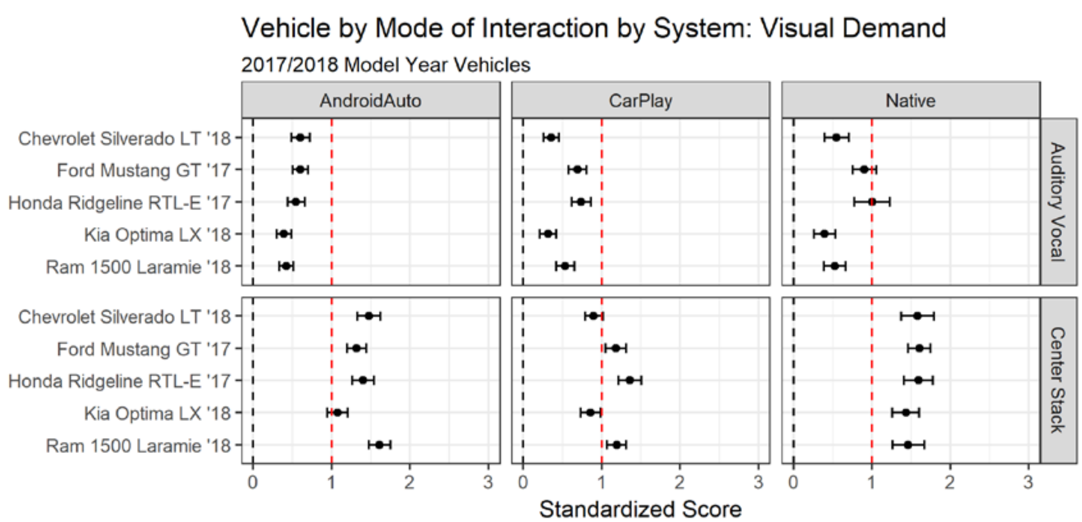

Figure 22. Visual demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(22) = 119.15$, $p < .01$).

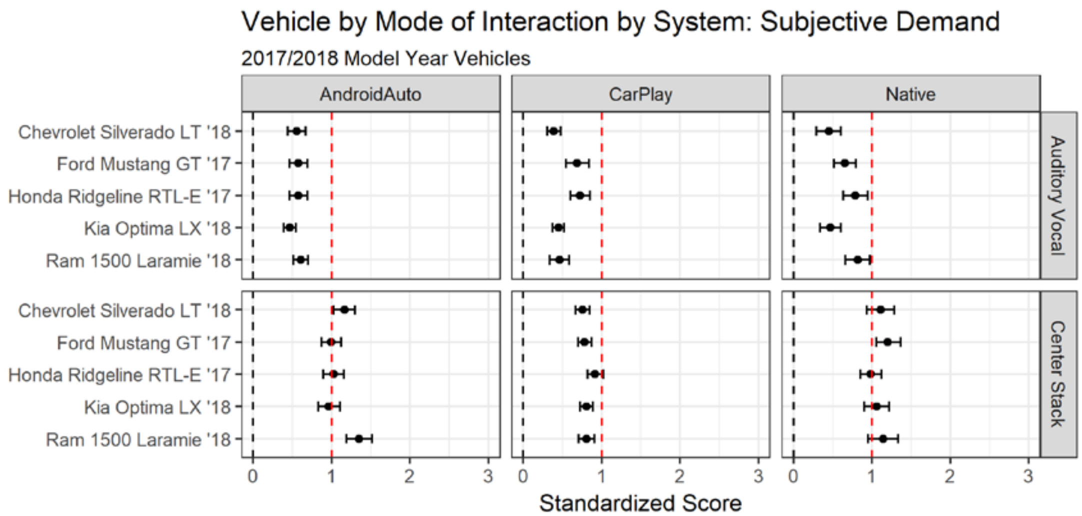

Figure 23. Subjective demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(22) = 120.47$, $p < .01$).

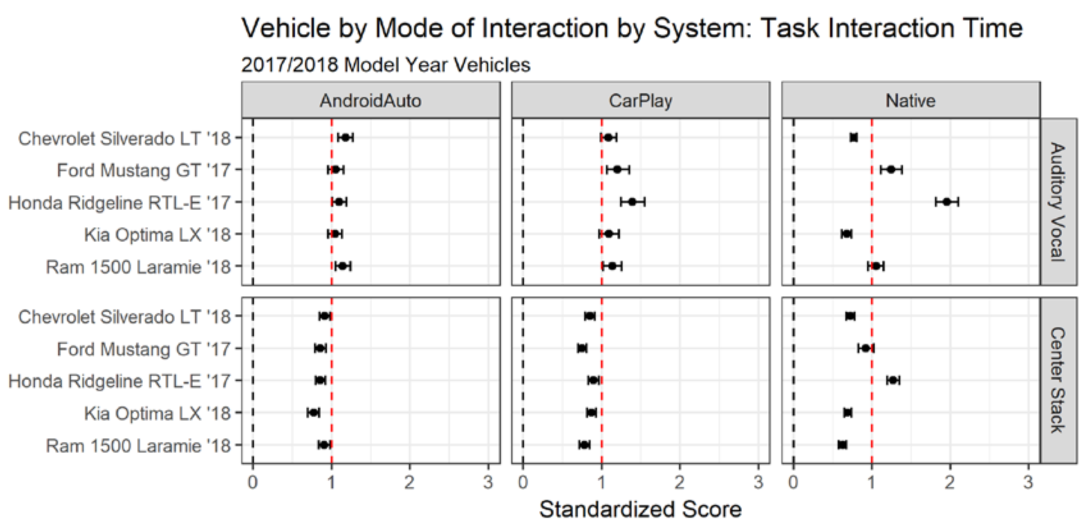

Figure 24. Task interaction time associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(22) = 277.13$, $p < .01$).

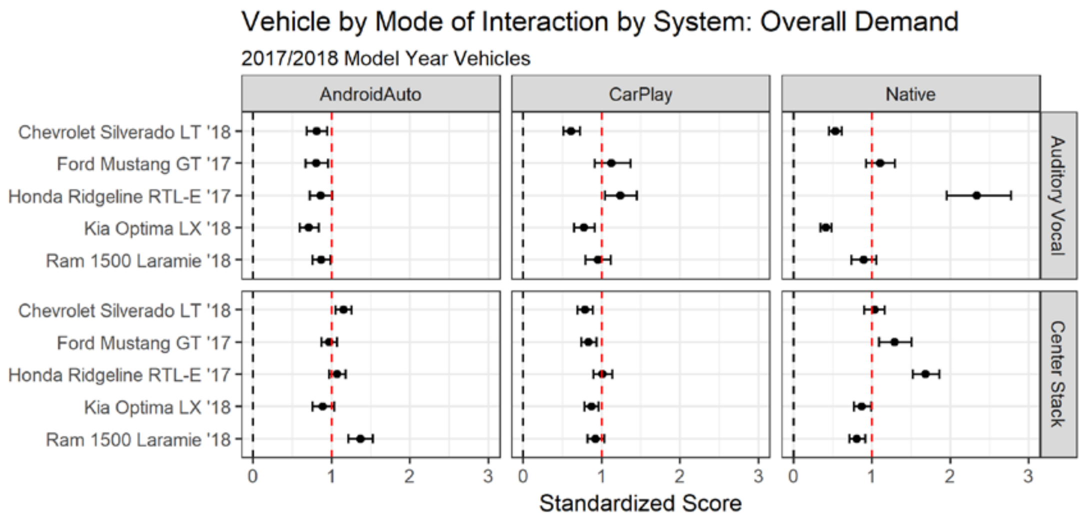

Figure 25. Overall demand associated with five different vehicles using the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(22) = 246.45$, $p < .01$).

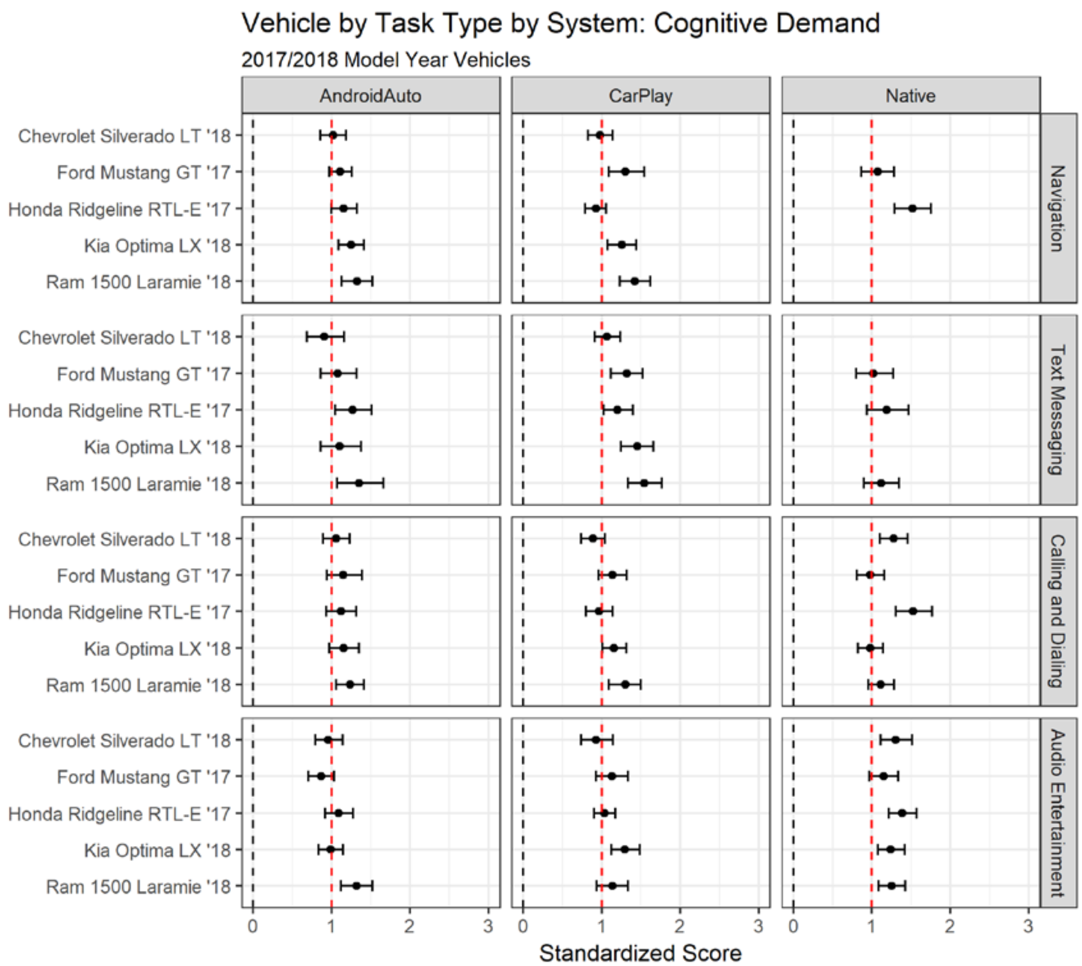

Figure 26. Cognitive demand associated with five different vehicles associated with navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Task by System interaction indicated that this interaction was significant ($X^2(45) = 102.07$, $p < .01$).

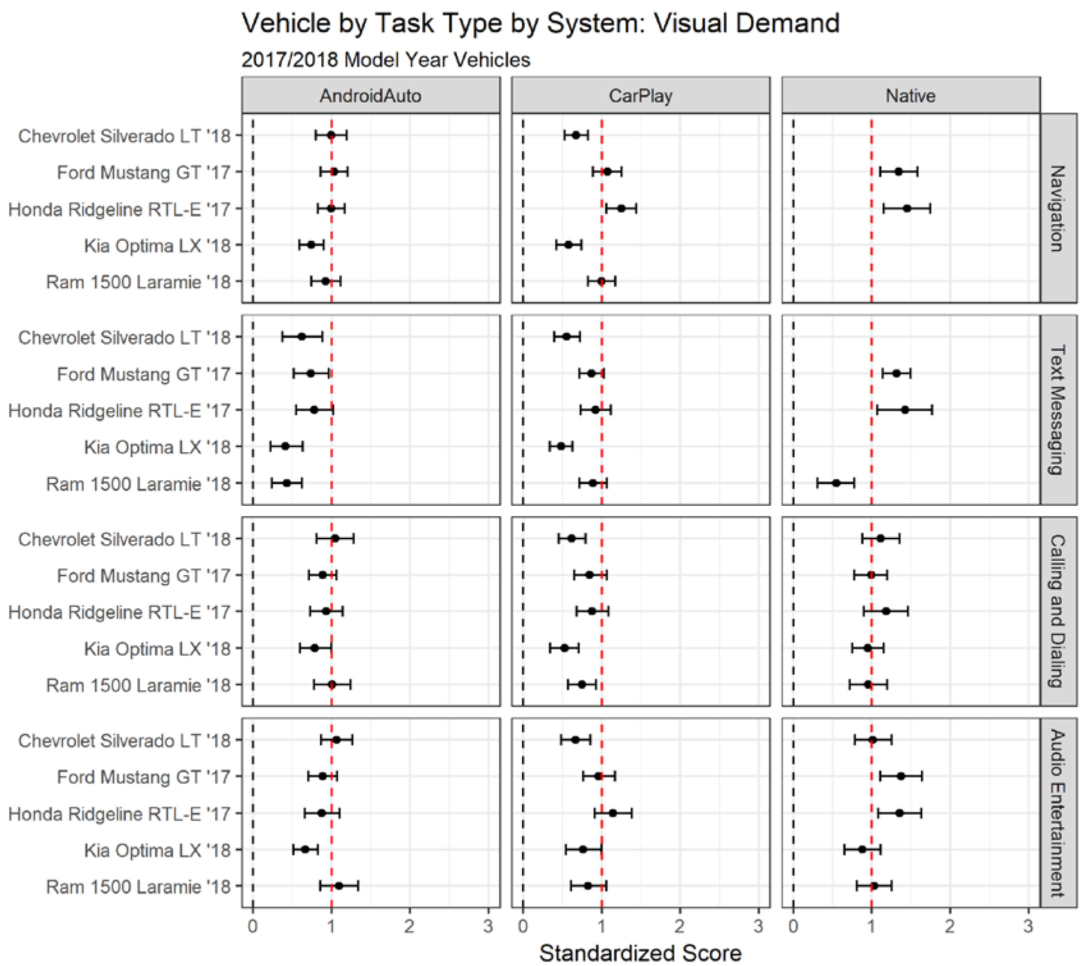

Figure 27. Visual demand associated with five different vehicles associated with navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Task by System interaction indicated that this interaction was significant ($X^2(45) = 100.39$, $p < .01$).

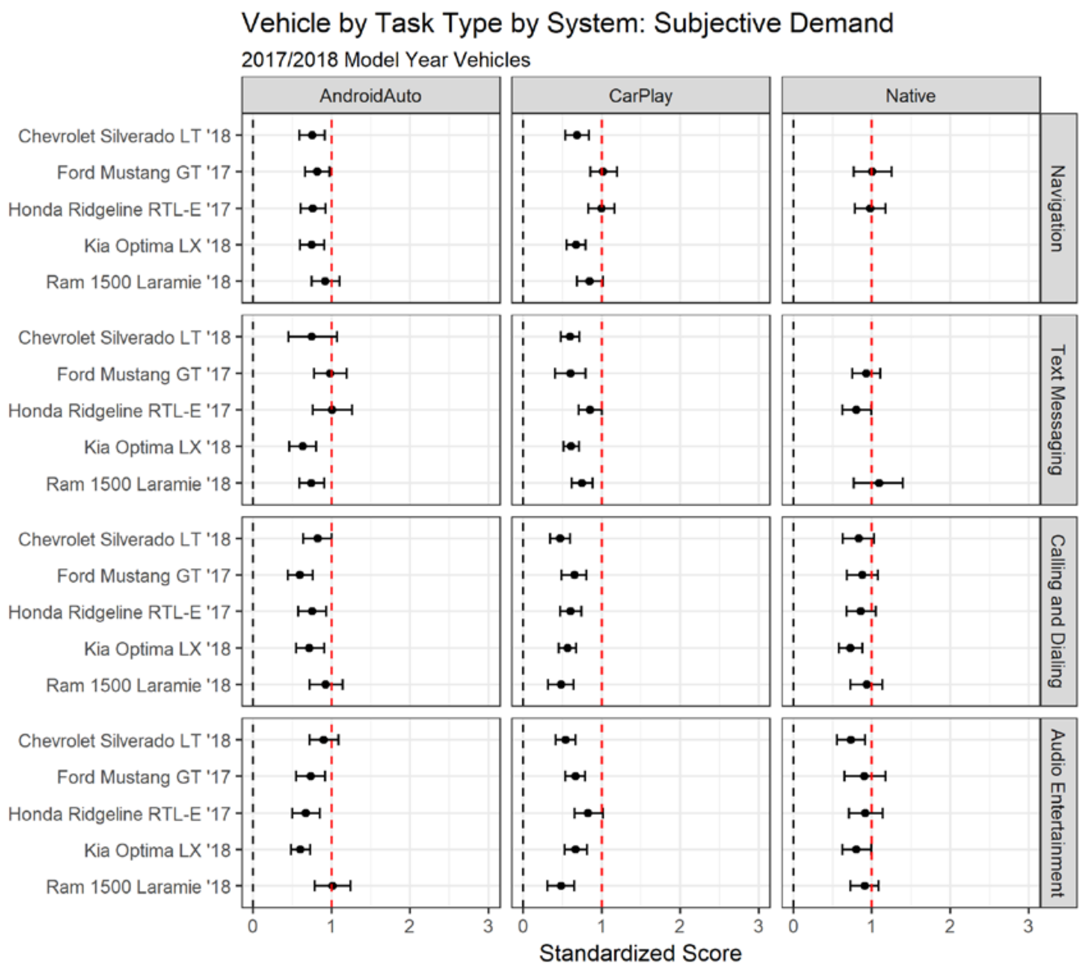

Figure 28. Subjective demand associated with five different vehicles associated with navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Task by System interaction indicated that this interaction was significant ($X^2(45) = 132.03$, $p < .01$).

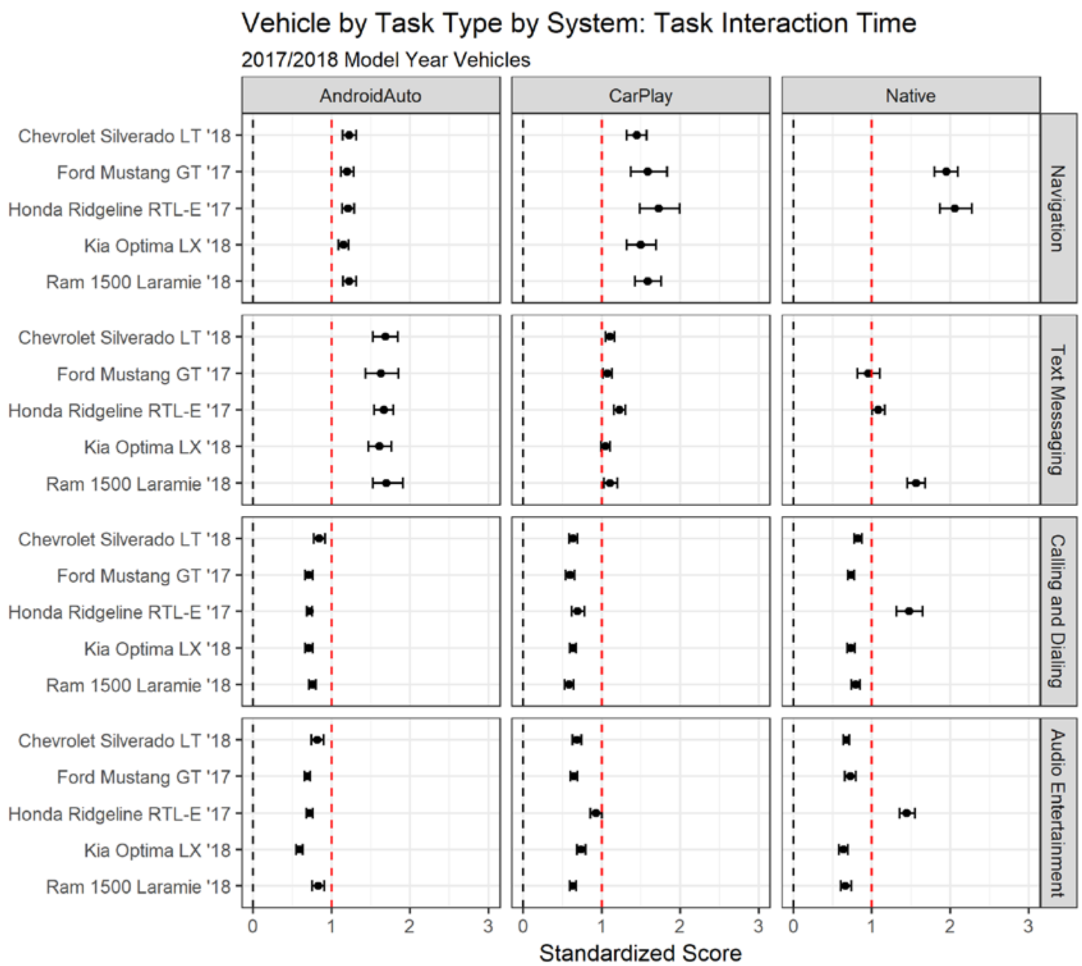

Figure 29. Task interaction time associated with five different vehicles associated with navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Task by System interaction indicated that this interaction was significant ($X^2(45) = 707.95$, $p < .01$).

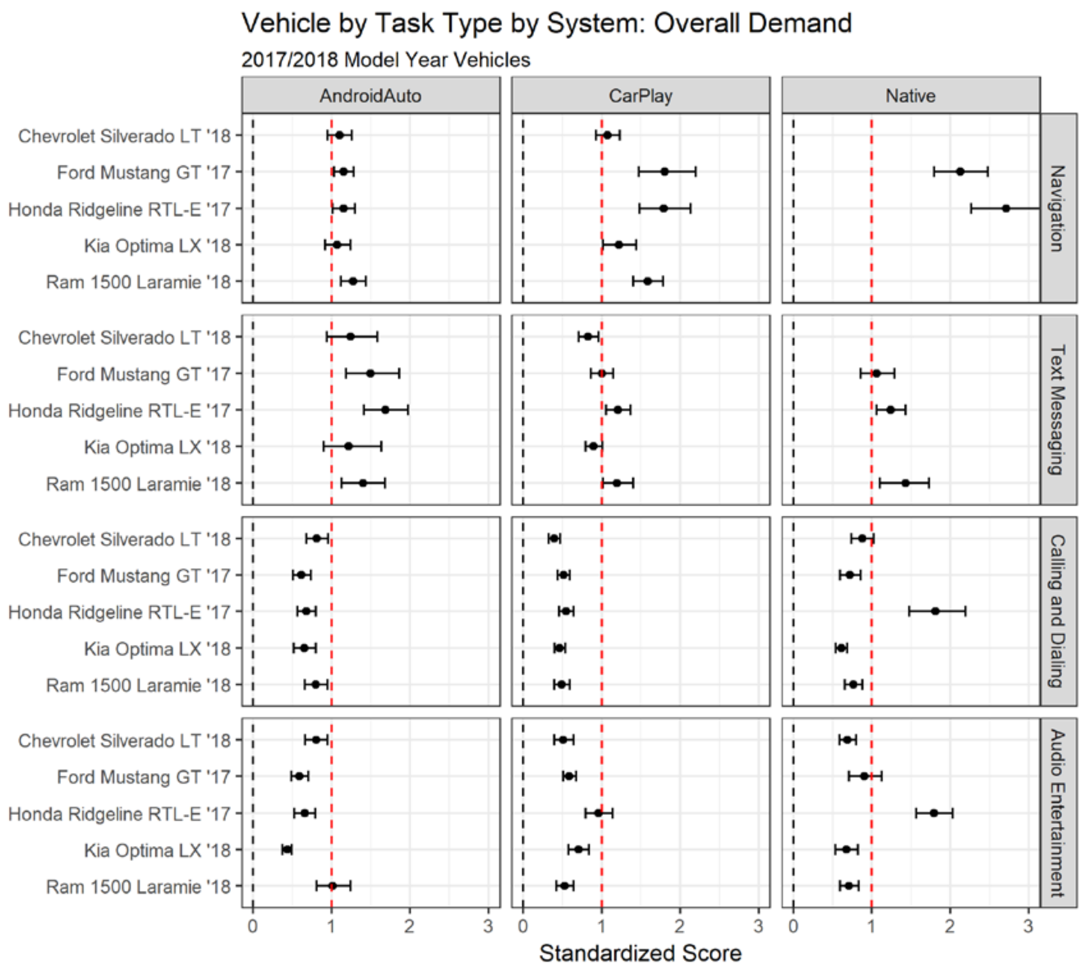

Figure 30. Overall demand associated with five different vehicles associated with navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Vehicle by Task by System interaction indicated that this interaction was significant ($X^2(45) = 376.38$, $p < .01$).

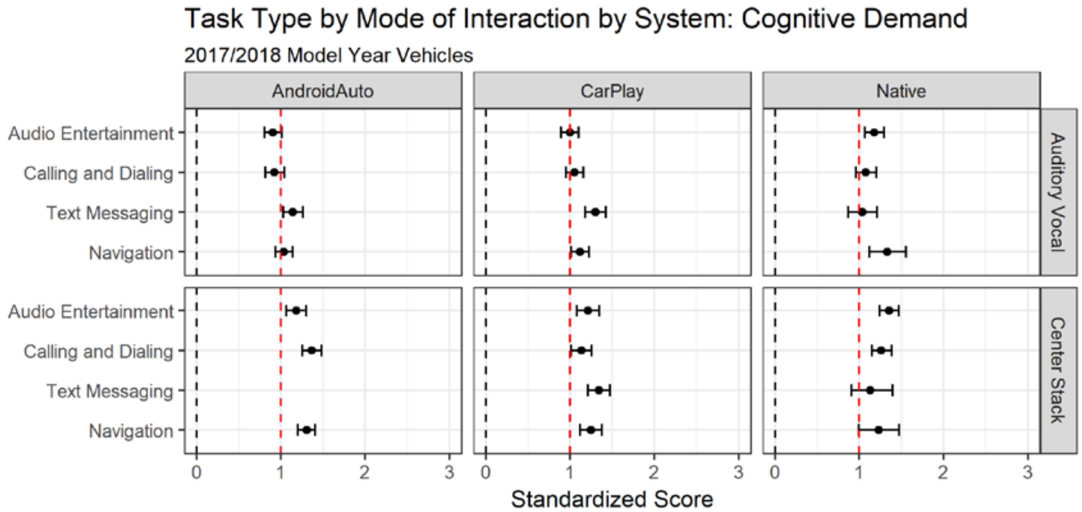

Figure 31. Cognitive demand associated with auditory/vocal and center stack interactions to support navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 1). The dashed black line represents single-task performance and the dashed red line represents the performance on the N-back task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Task by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(16) = 54.61$, $p < .01$).

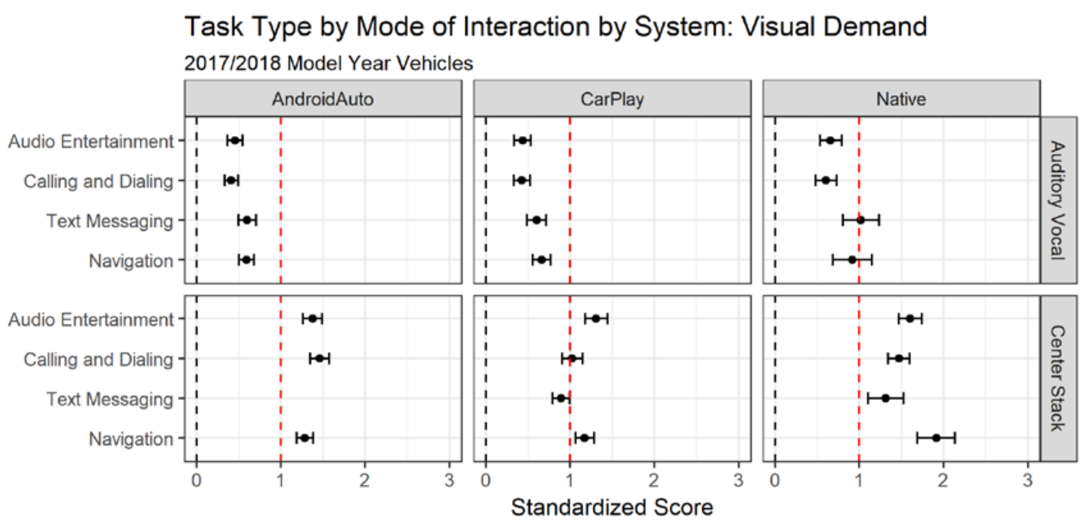

Figure 32. Visual demand associated with auditory/vocal and center stack interactions to support navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto system (Eq. 2). The dashed black line represents single-task performance and the dashed red line represents the performance on the SuRT task. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Task by Mode of interaction by Interface interaction indicated that this interaction was significant ($X^2(16) = 138.85$, $p < .01$).

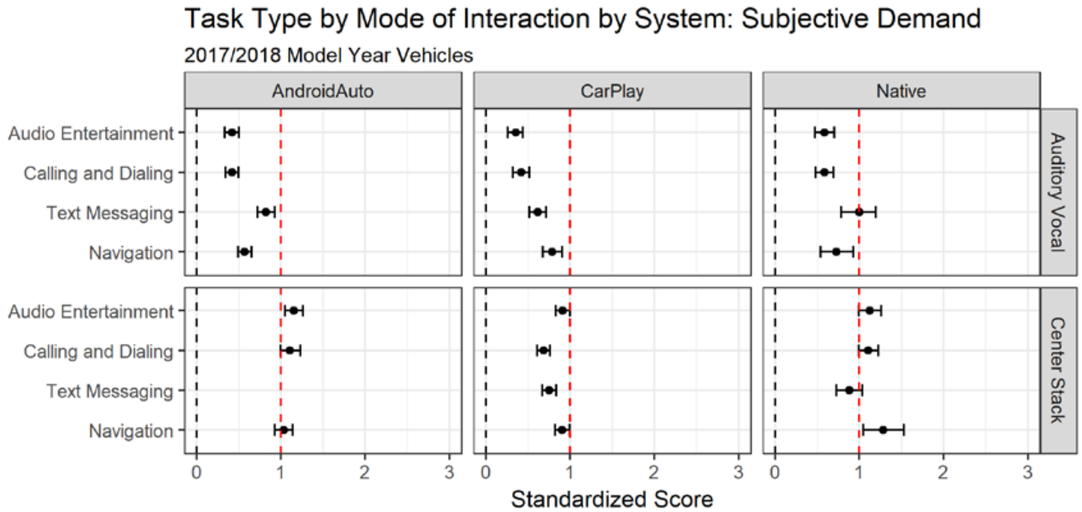

Figure 33. Subjective demand associated with auditory/vocal and center stack interactions to support navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 3). The dashed black line represents single-task performance and the dashed red line represents the average demand of the N-back and SuRT tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Task by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(16) = 160.53$, $p < .01$).

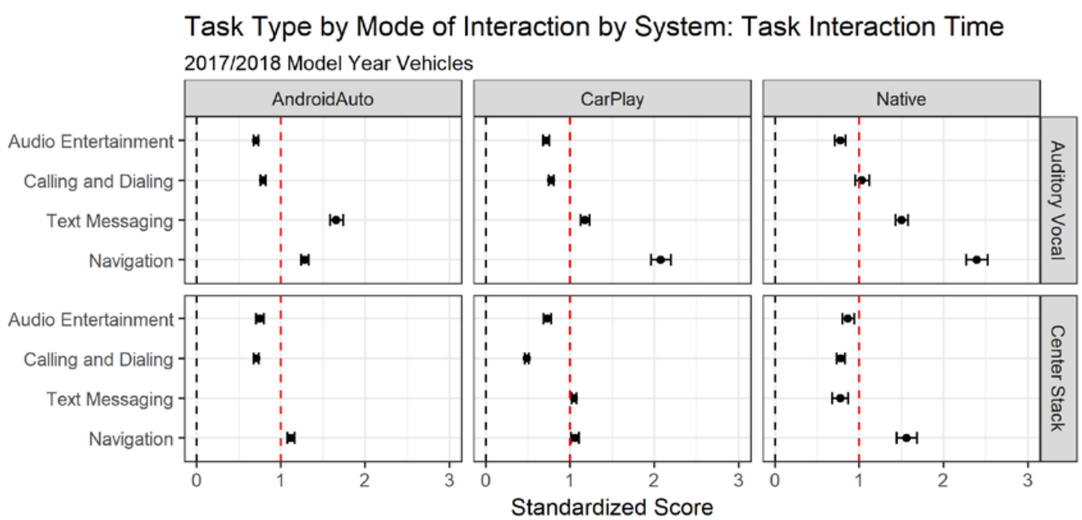

Figure 34. Task interaction time associated with auditory/vocal and center stack interactions to support navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 4). The dashed red line represents the 24-second task interaction referent. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Task by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(16) = 1199.71$, $p < .01$).

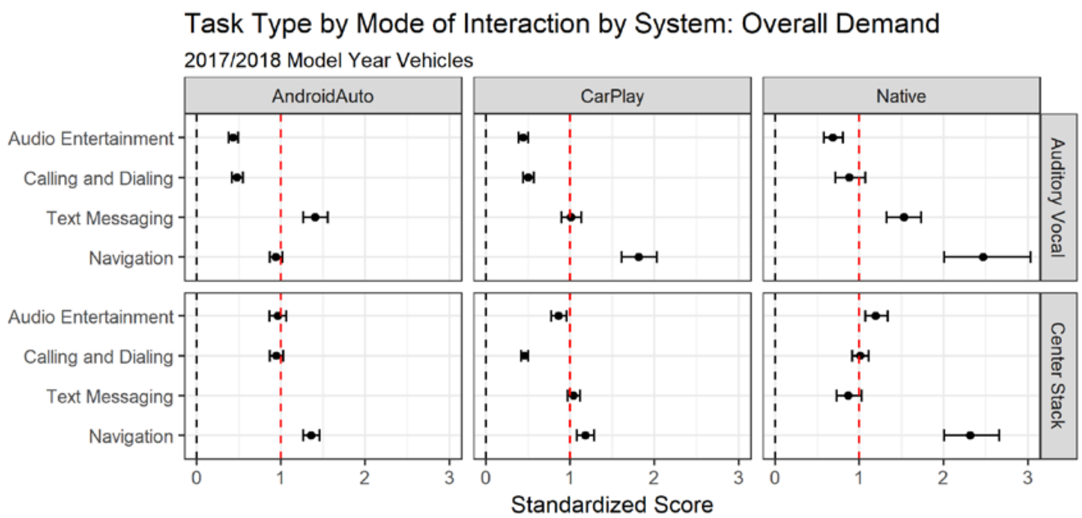

Figure 35. Overall demand associated with auditory/vocal and center stack interactions to support navigation, text messaging, calling and dialing, and audio entertainment using the native, CarPlay and Android Auto systems (Eq. 5). The dashed black line represents single-task performance and the dashed red line represents the high demand referent tasks. Error bars represent 95% confidence intervals. A comparison of linear mixed effects models with and without the Task by Mode of Interaction by System interaction indicated that this interaction was significant ($X^2(16) = 458.72$, $p < .01$).

Discussion

The current research compared the workload associated with using the in-vehicle information systems (IVIS) commonly available in five different automotive OEMs with that of CarPlay and Android Auto when used in the same vehicles. Both CarPlay and Android Auto are software platforms for the iPhone and Android smartphones, respectively, which allow the driver to pair their phone with the vehicle and perform many of the task types offered by the OEMs. These systems are often marketed as being easier to use than the native systems. We examined how the systems compared with the demand of the systems designed by the OEMs. Are the systems more or less demanding than what is provided by the OEMs? How do the systems compare with each other? Are there some tasks that are less demanding with one system than another (if so, why)? Finally, we examine whether these hybrid systems vary in demand with different vehicles.

The method of evaluation was identical to our earlier evaluation of IVIS interactions (Strayer, et al., 2017). The current study involved on-road testing of 24 participants in each configuration of five vehicles crossed with the three different infotainment systems: the native OEM system, CarPlay and Android Auto (i.e., each cell in the 5 X 3 factorial design had 24 participants). Replicating our prior research, we found that the task types differed in overall demand. The audio entertainment task type was found to be equivalent to the calling and dialing task type and both were significantly easier than the high demand referent (i.e., the red line). Text messaging and the navigation task types were harder than the high demand referent, with text messaging significantly less demanding than navigation.

A slightly different pattern to that reported in our earlier evaluation of IVIS interactions (e.g., Strayer et al., 2017) emerged when comparing the mode of interaction. Namely, our prior research found that auditory/vocal interactions were numerically (but not significantly) more demanding that center stack interactions. By contrast, in the current research the overall demand associated with auditory/vocal interactions was significantly lower than for center stack interactions. Comparison between the two studies found that the overall demand of center stack interactions did not differ, whereas the overall demand of auditory/vocal interactions was lower in the current study. In part, the reduced demand may be attributable to the superior auditory/vocal interface of Android Auto (as well as that of CarPlay, though to a lesser extent), when compared with the native OEM systems.

The analysis of current results found that the overall demand was significantly greater for the native OEM systems (which, on average, was higher than the high demand referent) than for CarPlay and Android Auto (see Figure 5). Additionally, the overall demand associated with using CarPlay and Android Auto did not differ and both systems were significantly below the high demand referent. This pattern addresses several of the research questions. In fact, the CarPlay and Android Auto systems are significantly less demanding than the native infotainment systems in the current set of vehicles. Although the overall demand did not differ between CarPlay and Android Auto, the systems have different strengths and weaknesses.

As shown in Figure 10, the mode of interaction influenced the overall workload. For CarPlay, the overall demand was nominally lower with center stack interactions than for auditory/vocal interactions. In contrast, for Android Auto, the overall demand was lower with auditory/vocal interactions than for center stack interactions. The strengths and weaknesses of each system traded off relative to each other in such a way that, when collapsed over mode of interaction, the overall demand of the interactions, as shown in Figure 5, did not differ. The native OEM systems were higher (significantly above the high demand referent) and more variable in demand for both modes of interaction.

As shown in Figure 15, the overall demand varied by task for CarPlay and Android Auto. For the audio entertainment task type, the overall demand was the same for the two systems and both were significantly less demanding to use than the native OEM systems. For the calling and dialing task type, CarPlay was significantly less demanding than Android Auto, which was less demanding than using the native OEM systems. For the text messaging task type, the overall demand was lower for CarPlay than for Android Auto with the native OEM systems falling in between (and not significantly different from) CarPlay and Android Auto. For the navigation task type, the overall demand of destination entry was significantly lower for Android Auto than for CarPlay, which was significantly less demanding than using the native OEM systems. As noted above, the strengths and weaknesses of each system traded off relative to each other in such a way that, when collapsed over task type, the overall demand of the interactions for CarPlay and Android Auto did not differ (Figure 5). With regard to our research question, this analysis shows that some task types are less demanding with one system than they are with another.

Figure 20 shows how the overall demand varies for different vehicles. Of special note is the tight cluster of demand scores for the different vehicles with Android Auto, a more variable set of ratings for CarPlay and the greatest variability for the native OEM systems. This pattern is seen both within a vehicle (e.g., smaller error bars for Android Auto) and across the different vehicles (i.e., both within and between, participant variability is lower with the hybrid systems than with the native OEM systems). Additionally, the differences between ratings are not due to hardware issues given that the same vehicle was used for testing. The differences also cannot be attributed to variability in the cellular network or the driving route, as these were held constant in our testing. Thus, the hybrid systems vary in demand when they are deployed in different vehicles. These differences are all the more striking given the greater functionality provided by CarPlay and Android Auto compared with the native OEM systems.

Finally, the higher-order interactions presented in Figures 25, 30 and 35 (as well as the intermediate figures that are used in their generation) are important to help identify precisely what sorts of interactions are most problematic with each type of interface. For example, consider the results from the navigation task type shown in Figure 35. The native OEM systems perform poorly whereas the overall demand of CarPlay and Android Auto is much lower. In fact, using the auditory/vocal interaction mode with Android Auto results in an overall demand that does not significantly differ from the high demand referent. Similar patterns of outcomes can be derived from the figures.

Android Auto and Apple CarPlay represent the latest technological approach to merging core mobile phone functionality with driving. The underlying design philosophy and architecture of each approach is largely similar. Both systems provide a vehicle-phone pairing that increases the capabilities of the infotainment systems using local and remote computing resources. Both systems provide access to a reconfigurable set of driver selected applications that are stored on the mobile phone. Both systems adapt the user experience to better fit the driving context. Our research suggests that this general approach may lead to reductions in driver workload even while providing expanded capabilities to drivers.

Limitations and Caveats

This research used the experimental method where participants were instructed to perform the tasks in a counterbalanced order. This method provides the ability to make causal statements regarding the workload associated with different systems. However, in real-world settings, drivers are free to perform the tasks if, when and where they choose. This complicates the relationship between driver workload as measured in experimental studies and crash risk. For example, motorists may attempt to self-regulate their non-driving activities, limiting them to periods where they perceive the risks to be lower. However, self-regulation depends upon drivers being aware of their performance and adjusting their behavior accordingly, an ability that is often limited by the same factors that caused them to be distracted in the first place (e.g., Sanbonmatsu et al., 2016).

This research was designed with the assumption that drivers will use technology that is available to them. No efforts were made to weigh results based on likelihood of use and all conditions were treated as equally important for the analysis. This research design leads to several important caveats that should be considered when interpreting the results. Principal among these is that the same set of tasks was not evaluated in each vehicle. Given the complexity of these real-world systems, a comparison of equal tasks with equivalent modes of interactions was not possible. A comparison between the few truly common tasks would have provided an unacceptably narrow and limited subset of the data. That said, the tested set of tasks should have favored the native interfaces as they supported fewer complex activities (e.g., navigation and texting were supported by two and three vehicles respectively). The finding that Android Auto and Apple CarPlay generally outperformed the native systems is, therefore, more remarkable.

The equal weighting of tested tasks also carries the implicit assumption that drivers will be equally likely to use each of the features and functions available in the vehicle. Given the unique ways in which these functions are being delivered by each of the systems, we saw no justifiable approach to determine whether certain tasks should be treated differently in the analyses than others. As technology improves it is very likely that users will change the way they interact with the systems. Usage patterns will likely evolve as better interfaces are developed and new functionality is introduced into vehicles. Thus, results presented in this report provide a snapshot of the overall demand profile of potential interactions that drivers may have with the vehicle and interface but not necessarily the actual demand that may be experienced by users of these systems on the roadway.

With respect to the benchmarks, we selected as high-demand referent tasks the N-back task (e.g., Mehler, Reimer, & Dusek, 2011) and SuRT (e.g., Engström & Markkula, 2007; Mattes, Föhl, & Schindhelm, 2007) and adopted a 24-second rule for dynamic task interaction time. The 24-second task interaction referent was derived based on the project team’s interpretation of the NHTSA visual/manual guidelines (NHTSA, 2013). One may question whether these referents are reasonable. For example, if the referent tasks were too easy (or hard), then the absolute ratings would be an overestimate (or underestimate) of the true demand. However, it is important to note that the relative ratings will be insensitive to the absolute demand of the referent tasks, so long as they are performed in a consistent fashion in a counterbalanced order across participants.

Summary

CarPlay and Android Auto provided more functionality and resulted in lower levels of workload than the native OEM infotainment systems. That said, both incurred moderately high levels of demand, thus providing opportunities to improve the user experience. For example, CarPlay had lower overall demand than Android Auto for sending text messages and Android Auto had lower overall demand than CarPlay for destination entry to support navigation. In light of the current results, both systems can be improved toward more optimal user experiences.

References

Angell, L. S., Auflick, J., Austria, P. A., Kochhar, D. S., Tijerina, L., Biever, W., ... & Kiger, S. (2006). *Driver Workload Metrics Task 2 Final Report* (Rep. HS-810 635). Washington, D.C.: National Highway Traffic Safety Administration.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Burns, P., Harbluk, J., Foley, J. P., & Angell, L. (2010). The importance of task duration and related measures in assessing the distraction potential of in-vehicle tasks. *Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. doi:10.1145/1969773.1969776.

Castro, S., Cooper, J., & Strayer, D. (2016). Validating two assessment strategies for visual and cognitive load in a simulated driving task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*, 1899-1903.

Cooper, J. M., Castro, S. C., & Strayer, D. L. (2016). Extending the Detection Response Task to simultaneously measure cognitive and visual task demands. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*, 1962-1966.

Davison, A. C., Hinkley, D. V, & Young, G. A. (2003). Recent developments in bootstrap methodology. *Statistical Science*, *18*(2), 141–157.

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 97–120.

Engström, J. & Markkula, G. (2007). Effects of visual and cognitive distraction on lane change test performance. *Proceedings of the Fourth International Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.* Iowa City, IA: University of Iowa.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*, 323-343.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346-363.

International Organization for Standardization (2012). *Road Vehicles – Ergonomic Aspects of Transport Information and Control Systems – Calibration Tasks for Methods which Assess Driver Demand Due to the Use of In-Vehicle Systems* (Rep. ISO TS 14198). Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization (2015). *Road Vehicles – Transport Information and Control Systems – Detection Response Task (DRT) for Assessing Attentional Effects of Cognitive Load in Driving* (Rep. ISO 17488). Geneva, Switzerland: International Organization for Standardization.

Ito, H., Atsumi, B., Uno, H., & Akamatsu, M. (2001). Visual distraction while driving: Trends in research and standardization. *IATSS Research, 25*(2), 20-28.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Kidd, D. G., Dobres, J., Reagan, I., Mehler, B., & Reimer, B. (2017). Considering visual-manual tasks performed during highway driving in the context of two different sets of guidelines for embedded in-vehicle electronic systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, *47*, 23-33. doi:10.1016/j.trf.2017.04.002.

Little, T. D. & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives, 7*, 199-204.

Mattes, S., Föhl, U., & Schindhelm, R. (2007). *Empirical Comparison of Methods for Off-line Workload Measurement* (AIDE Deliverable 2.2.7, EU Project IST-1-507674-IP). Gothenburg, Sweden: Adaptive Integrated Driver-vehicle InterfacE.

McWilliams, T., Reimer, B., Mehler, B., Dobres, J., & Coughlin, J. F. (2015). Effects of age and smartphone experience on driver behavior during address entry. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. doi:10.1145/2799250.2799275.

Mehler, B., Kidd, D., Reimer, B., Reagan, I., Dobres, J., & McCartt, A. (2015). Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics*, *139*, 1-24. doi:10.1080/00140139.2015.1081412.

Mehler, B., Reimer, B., & Dusek, J. (2011). *MIT AgeLab Delayed Digit Recall Task (n-back)* (White paper). Cambridge, MA: MIT AgeLab.

National Highway Traffic Safety Administration (2013). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices* (Federal Register Vol.78, No. 81). Washington, D.C.: National Highway Traffic Safety Administration.

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/.

Reimer, B., Mehler, B., Dobres, J., McAnulty, H., Mehler, A., Munger, D., & Rumpold, A. (2014). Effects of an “Expert Mode” voice command system on task performance, glance behavior & driver physiology. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. doi:10.1145/2667317.2667320.

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, *43*, 1771-1781.

Regan, M. A. & Strayer, D. L. (2014). Towards an understanding of driver inattention: Taxonomy and theory. *Annals of Advances in Automotive Medicine*, *58*, 5-13.

Society of Automotive Engineers (2015). *Operational Definitions of Driving Performance Measures and Statistics* (No. J2944). Warrendale, PA: Society of Automotive Engineers.

Sanbonmatsu, D. M., Strayer, D. L., Biondi, F., Behrends, A. A., & Moore, S. M. (2016). Cell phone use diminishes self-awareness of impaired driving. *Psychonomic Bulletin & Review, 23*, 617-623.

Schmidt, R. A. & Lee, T. D. (2005). *Motor Control and Learning: A Behavioral Emphasis* (Vol. 4). Champaign, IL: Human kinetics.

Shutko, J. & Tijerina, L. (2006). Eye glance behavior and lane exceedences during driver distraction. Presentation given at *Driver Metrics Workshop*, Ottawa, Canada. Available at http://ppc.uiowa.edu/drivermetricsworkshop/.

Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D., & Biondi, F. (2017). *Visual and Cognitive Demands of Using In-Vehicle Information Systems*. Washington, D.C.: AAA Foundation for Traffic Safety.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. In B. Ross (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory (Vol 54)* (pp. 29-58). San Diego, CA: Elsevier Academic Press.

Tsimhoni, O., Yoo, H., & Green, P. (1999). *Effects of Visual Demand and In-Vehicle Task Complexity on Driving and Task Performance as Assessed by Visual Occlusion* (Technical Report UMTRI-99-37). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Wickens, C. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449-455. doi:10.1518/001872008X288394.

Young, R. A., Aryal, B. J., Muresan, M., Ding, X., Oja, S., & Simpson, S. N. (2005). Road-to-lab: Validation of the static load test for predicting on-road driving performance while using advanced in-vehicle information and communication devices. *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Iowa City, IA: University of Iowa.

Zhang, Y., Angell, L., Pala, S., & Shimonomoto, I. (2015). Bench-marking drivers' visual and cognitive demands: A feasibility study. *SAE International Journal of Passenger Cars—Mechanical Systems, 8*(2), 584-593. doi:10.4271/2015-01-1389.

Appendix 1

A. Audio Entertainment Tasks: Restricted Content

- Restricted iPod contents (songs, artists, albums, genres)
- *Experimenter Note*: All list content must be played before start of experimental condition because only the 10 most recent selections are available while the vehicle is in motion.

1. Choose a **jazz song**	**11.** Switch the music to **artist Pantera**
2. Listen to the **song "99 Red Balloons"**	**12.** Play the **alternative genre**
3. You want to hear a song by the **band Nirvana**	**13.** Change the **genre** to **reggae**
4. Play a song in the **metal genre**	**14.** You want to hear a song by the **artist Hunter Hayes**
5. Let's hear the **song "I'm Gonna Be (500 Miles)"**	**15.** Let's listen to the **rock genre**
6. Play **album "Storyline"**	**16.** Change the music to the **song "If It Means a Lot to You"**
7. Play the song **"Smells Like Teen Spirit"**	**17.** The **album "Dream Your Life Away"** is what we should hear
8. Play the **album "Homesick"**	**18.** The **artist The Proclaimers** is what we should listen to
9. You want to play **Johnny Cash songs**	
10. "Riptide" is a **song** you want to play	

B. Audio Entertainment Tasks: Full Content

- Full iPod contents (songs, artists, albums, genres)
- Used when system did not restrict scrolling within category lists

1. Choose a **jazz song**

2. Listen to the **song “99 Red Balloons”**

3. The **band Nirvana** is what you want to hear

4. Turn on the **metal genre**

5. Let’s hear the **song “I’m Gonna Be (500 Miles)”**

6. Play **album “Storyline”**

7. Play a song by the **artist Eminem**

8. Play the **album “Homesick”**

9. You want **Johnny Cash songs** to play

10. “Riptide” is a **song** you want to play

11. Switch the music to **artist Louis Armstrong**

12. Play the **alternative genre**

13. Change the **genre** to **reggae**

14. You want to hear a song by the **artist Hunter Hayes**

15. Change the music to the **song “Three Little Birds”**

16. Let’s listen to the **rock genre**

17. The **album “Let it Be”** is what we should hear

18. The **artist Florida Georgia Line** is what we should listen to

19. Play **song, “What a Wonderful World”**

20. Let’s listen to something by **artist Katy Perry**

21. Jam out to the **artist Adele**

22. You want to hear **song “Super Bass”**

23. Let’s hear the **album “Rise”**

24. The **artist Vance Joy** is who you want to hear

25. Let’s listen to the **hip-hop genre**

26. Play **song “Somebody Else”**

27. Artist Slip Knot is what we should listen to

28. Play **album, “Kind of Blue**”

29. Get down to some **country** music

30. Journey the **artist** is who you want to hear

C. Audio Entertainment Tasks: Song Titles Only

- Restricted iPod contents (songs, artists, albums, genres)
- *Experimenter Note*: All list content must be played before start of condition because only the 10 most recent selections are available while the vehicle is in motion.

1. Choose a **song**	**11.** Switch to the **song "Let it Be"**
2. Listen to the song **"99 Red Balloons"**	**12.** Play the song **"Come Together"**
3. The **song "Not Afraid"** is what you want to hear	**13.** Change the music to the **song "Three Little Birds"**
4. Turn on **"Mess Around" the song**	**14.** Let's play the **song, "The Funeral"**
5. Let's hear the **song "I'm Gonna Be (500 Miles)"**	**15.** The **song "Super Bass"** is what you want to hear
6. Phone play **album "Storyline"**	**16.** You want to listen to **"Walk"** the **song**
7. Play the **song "Don't Stop Believin"**	**17. Song "Sorry"** is what you want to listen to
8. Play the **song "I Can See Clearly Now"**	**18.** Play **song "Somebody Else"**
9. You want **"If It Means a Lot To You" song** to play	**19.** Jam out to the **song, "What a Wonderful World"**
10. "Riptide" is a **song** you want to play	**20.** Change the music to, **"Smells Like Teen Spirit"**

D. **Audio Entertainment Tasks: Radio and iPod Content**

- Radio frequency tuning
- iPod content (songs, artists, albums, genres)

1. Choose a **jazz song** from the iPad	**14. AM 1160**
2. Play **1020 AM**	**15. Play a song by the artist Eminem**
3. Tune the radio to **98.5 FM**	**16.** Play the **album "Homesick"**
4. Listen to the song **"99 Red Balloons"**	**17. 90.1**
5. The **band Nirvana** is what you want to hear	**18.** You want **Johnny Cash songs** to play
6. Change the radio to your favorite **FM station**	**19.** Radio **1630**
7. Turn on the **rock genre**	**20. "Riptide" is a song you want to play** from the iPod
8. Let's hear the **song "I'm Gonna Be (500 Miles)"**	**21.** Switch the iPod to **artist Louis Armstrong**
9. You want to hear one of your favorite **AM stations**	**22.** Play the **alternative genre**
10. Tune **AM 1540**	**23.** Change the **genre** to **reggae**
11. 89.1	**24.** Radio tune to **97.1 FM**
12. Tune to **1240**	**25.** You want to hear a song by the **artist Hunter Hayes**
13. iPod play **album "Stomping Ground"**	**26.** Change the music to the **song "Three Little Birds"**
	27. Listen to **FM 99.5**

E. Audio Entertainment Tasks: Radio and iPod as Source

- Radio frequency tuning
- iPod as input source (no content access)

1. Choose a song from the **iPad** **2.** Play **1020 AM** **3.** Tune the radio to **98.3 FM** **4.** Listen to the song on the **iPod** **5.** Change the radio to your favorite **FM station** **6.** Turn on the **iPod** **7.** Let's hear a song on the **iPad** **8.** You want to hear one of your favorite **AM stations** **9.** Tune **AM 1540** **10. 89.1** **11.** Tune to **1240**	**12.** iPod source **13. AM 1160** **14.** Play a song via **iPad** **15. 90.1** **16.** You want **iPod** songs to play **17.** Radio **1630** **18.** You want to play music from the **iPod** **19.** Radio tune to **97.1 FM** **20.** You want to hear a song by your favorite artist on the **iPod** **21.** Listen to **FM 99.5**

F. Calling & Dialing Tasks: Contacts Only

- Participant calls contacts (cell phone, work)
- Task is complete once call has been successfully ended

1. **Jack Olsen** would like you to call him on his **cell phone**
2. **Willow Brooks**
3. Try to reach **Brad Peterson**
4. Ring **Felicity Gomez's cell**
5. You missed a call from **Oliver Reed**
6. **Violet Wheeler** is waiting to hear back from you on her **mobile**
7. Give **Phil Potter** a call back
8. Try **Helen Harold** on her **mobile number**
9. **Bethany Swan, cell phone**
10. Telephone **Jennifer Long**
11. You need to talk to **Willow Brooks**
12. Dial **Brad Peterson**
13. Call **Jack Olsen mobile** back
14. You need to call **Jennifer Long**
15. Place a call to **Helen Harold**
16. You can't reach **Bethany Swan**. Call them again.
17. You need to reach **Oliver Reed**
18. Telephone **Phil Potter**
19. Dial **Felicity Gomez, mobile**
20. Give **Violet Wheeler** a call

G. Calling & Dialing Tasks: Contacts and Dialing

- Participant calls contacts (cell phone, work)
- Dials numbers (participant's own phone number, 801-555-1234)
- Task is complete once call has been successfully ended

1. **Jack Olsen** would like you to call him on his **cell phone**	**11.** **Violet Wheeler** is waiting to hear back from you on her **mobile**
2. You need to call **8"OH"1-555-1234**	**12.** Dial your **own number**
3. **Willow Brooks**	**13.** Give **Phil Potter** a call back at work
4. Try to reach **Brad Peterson**	**14.** Give **8"ZERO"1-555-1234** a call
5. Enter **8"ZERO"1-555-1234**	**15.** Try **Helen Harold** on her **business number**
6. You can't find **your phone**. Call it to find it.	**16.** Call your **own phone**
7. Ring **Felicity Gomez's office**	**17.** **8"OH"1-555-1234**
8. Enter your **own number**	**18.** **Bethany Swan, cell phone**
9. You missed a call from **Oliver Reed**	**19.** Telephone **Jennifer Long**
10. Telephone **8"OH"1-555-1234**	**20.** You need to talk to **Yolanda Chavez**

H. Navigation Tasks

- Participant sets the destination to a point of interest that best fits the task goal
- Participant cancels the route before the task is considered to be complete

1. **(Gas)** Fill up at the closest gas station. 2. **(Coffee)** Grab yourself a cup of coffee from Cafe on 1st. 3. **(Parking)** You are going shopping. Let’s find parking. 4. **(Grocery Store)** Let’s grab some items from Whole Foods. 5. **(Restaurant)** You are hungry for some cheesecake; go to a place to satisfy your craving. 6. **(Gas)** Let’s find a Tesoro to gas up the car at.	7. **(Restaurant)** A burger place is where we should go to eat. 8. **(Coffee)** Let’s get some coffee at a nearby coffee shop. 9. **(Grocery Store)** Let's pick up some items at the nearby grocery store. 10. **(Restaurant)** Pizza sounds great right now; navigate to a pizza place. 11. **(Parking)** You want to go walking downtown; let’s find some parking nearby.

I. **Navigation Tasks (with other locations)**

- Participant sets the destination to a point of interest that best fits the task goal
- Participant cancels the route before the task is considered to be complete

1. (Gas) Fill up at the closest gas station. **2. (Library)** Your library book is overdue. Let's return it at the closest library. **3. (Italian Restaurant)** You're headed out for some Italian food at nearby restaurant. **4. (Coffee shop)** Grab yourself a cup of coffee from the closest Starbucks. **5. (Grocery store)** You need some items from Whole Foods. **6. (ATM\bank)** You need to get cash from a Wells Fargo bank. **7. (Mexican Restaurant)** Find a Mexican restaurant nearest you.	**8. (Hospital)** Go visit your friend at the LDS Hospital. **9. (Chinese Restaurant)** You're craving food from Panda Express. **10. (Movie theater)** You're on your way to see a movie at the nearby theater. **11. (Hotel/Motel)** Drive to the nearest lodging to stay the night. **12. (Post office)** You have a package to drop off at the closest post office. **13. (Museum)** Go check out the new exhibit at the Utah Museum of Natural History. **14. (Shopping Center)** Go pick out some new clothes at a nearby shopping mall.

J. SMS Tasks: Send Only

- Participant sends a new text message to a phone contact in response to the given scenario
- Task is complete once message has been sent
- *Experimenter Note*: Researcher must clear inbox of any messages not from the contacts in the list below

1. Let **Cam Whitman** know you're going to be late.
2. **Kaley Mackenzie** is asking if you want to go to the movies tonight.
3. **Dilbert Pugsley** wants to go dancing tonight.
4. **Maggie Carter** texted you a funny joke.
5. Text **Rachel Gatsby** to ask for directions.
6. Ask **Dilbert Pugsley** where they are.
7. Tell **Andy Cameron** you're too busy driving to text right now.
8. Tell **Dilbert Pugsley** to text you.
9. **Maggie Carter** texted you another a silly dad joke.
10. **Cam Whitman** has big news and is wondering if you can talk right now.
11. **Cam Whitman** dropped off your favorite cookies at your house.
12. **Kaley Mackenzie** is wondering where you are.
13. **Maggie Carter** can pick you up from the airport next week.
14. Tell **Rachel Gatsby** to call you from work
15. Tell **Dilbert Pugsley** you're too busy driving to call them right now.
16. **Andy Cameron** wants to know if they can copy your homework.
17. **Cam Whitman** wants to know why you're not at the restaurant yet.
18. **Maggie Carter** says she will clean your car for you tonight.

K. SMS Tasks: Send Only

- Participant sends a new text message in response to the given scenario to a phone contact
- Task is complete once message has been sent

1. Let **Hugo Grant's office** know you're going to be late.	**11. Zoe Ferris** dropped off your favorite cookies at your house.
2. Hunter Bowman is asking if you want to go to the movies tonight.	**12. Isabelle Morales** is wondering where you are.
3. Eve Remington wants to go dancing tonight.	**13. Natalie Ling** can pick you up from the airport next week.
4. Milly Jung texted you a funny joke.	**14.** Tell **Jack Olsen** to call you from **work.**
5. Text **Kevin Malcome** to ask for directions.	**15.** Tell **Willow Brooks** you are too busy driving to call them right now.
6. Ask **Quinn Brown (cell)** where they are.	**16. Francis Baker** wants to know if they can copy your homework.
7. Tell **Paige Green** you're too busy driving to text right now.	**17. Milo Santiago** wants to know why you're not at the restaurant yet.
8. Tell **Landon Carter** to text you.	**18. Gretchen Warner** says they will clean your car for you tonight.
9. Vince Hancock texted you a silly dad joke.	
10. Brad Peterson has big news and is wondering if you can talk right now.	

L. SMS Tasks: Send Only (Ford Mustang GT Native System)

- Participant sends a new text message in response to the given scenario to a phone contact
- Task is complete once message has been sent

1. Let **Lucas Forester**[5] know you're going to be late.	**11. Lucas Forester** dropped off your favorite cookies at your house.
2. Lucas Forester is asking if you want to go to the movies tonight.	**12. Lucas Forester** is wondering where you are.
3. Lucas Forester wants to go dancing tonight.	**13. Lucas Forester** can pick you up from the airport next week.
4. Lucas Forester texted you a funny joke.	**14.** Tell **Lucas Forester** to call you
5. Text **Lucas Forester** to ask for directions.	**15.** Tell **Lucas Forester** you're too busy driving to call them right now.
6. Ask **Lucas Forester** where they are.	**16. Lucas Forester** wants to know if they can copy your homework.
7. Tell **Lucas Forester** you're too busy driving to text right now.	**17. Lucas Forester** wants to know why you're not at the restaurant yet.
8. Tell **Lucas Forester** to text you.	**18. Lucas Forester** says they will clean your car for you tonight.
9. Lucas Forester texted you a silly dad joke.	
10. Lucas Forester has big news and is wondering if you can talk right now.	

[5] The same name was used for the different tasks because for voice conditions using SYNC3, the system would only reply to the most recent message in the inbox (from Lucas Forester).

M. SMS Tasks: Read Only

- Participant has system read out text messages
- Task is complete once message has been selected, not once message is done being read aloud

1. Read out the message from **Cam Whitman**
2. Read out the text from **Andy Cameron**
3. What did **Amelia Kidder** send you?
4. **Maggie Carter** just messaged you.
5. What did **Rachel Gatsby** say?
6. Find a message from **Andy Cameron**
7. **Scarlett Miles** sent you a new text
8. Read the text from **Amelia Kidder**
9. What did **Maggie Carter** send you?
10. Read the text from **Cam Whitman**
11. New message from **Lucas Forester**
12. What did **Cam Whitman** send you?
13. Read out the message from **Maggie Carter**
14. What did **Scarlett Miles** send you?
15. What does the text from **Maggie Carter** say?
16. What did **Rachel Gatsby** send you?
17. What does the message from **Scarlett Miles** say?
18. New message from **Andy Cameron**

N. SMS Tasks: Read & Send

- Participant reads and responds to text messages with system-specific predetermined messages
- Task is complete once message has been sent

1. Read out the message from **Cam Whitman**. Please respond.	**9.** Read and respond to the text from **Cam Whitman**
2. Read and reply to the text from **Andy Cameron**	**10.** What did **Cam Whitman** send you? Answer him.
3. What did **Amelia Kidder** send you? Send your answer.	**11.** Read out the message from **Maggie Carter.** Send your reply.
4. **Maggie Carter** just messaged you. What should you send back?	**12.** What did **Scarlett Miles** send you? Text her back.
5. Find a message from **Andy Cameron**. Reply.	**13.** How do you respond to the text from **Maggie Carter**?
6. **Scarlett Miles** sent you a new text. Send something back.	**14.** You need to read and reply to **Rachel Gatsby's** message.
7. Read the text from **Amelia Kidder** and respond to it.	**15.** What does the message from **Scarlett Miles** say? Respond.
8. What did **Maggie Carter** send you? Send a text back.	**16.** Reply to the new message from **Andy Cameron.**

The similarities and differences between each version of Android Auto and Apple CarPlay in each vehicle tested are outlined below. Variations of Android Auto are described in Sections A-C. Variations of Apple CarPlay are described in Sections D-E. Within each section, the two modes of interaction (center stack and voice command system) are discussed separately. Original Equipment Manufacturers (OEMs) determine which features are available while the vehicle is in motion. Therefore, unique characteristics of each OEM implementation are further explained.

Vehicles tested in simultaneous time frames with the same phone operating system and application version (if applicable) were grouped together. Since application and operating system updates are frequent and unpredictable, the auto-update feature on all phones was turned off. This ensured all participants interacted with the same application version and operating system within each vehicle. Applications and phone operating systems were not updated until data collection from all participants for the individual vehicle was completed, at which point the applications and operating systems were checked for updates that would be installed when available for the next group of vehicles being tested.

A. ANDROID AUTO – 2017 Ford Mustang GT and 2017 Honda Ridgeline RTL-E

Tested with the following software:

- Android Phone Operating System (OS) 7.1.2
- Android Auto App Version v2.6.573463

Center Stack Comparisons

1. *Center Stack Structure*

The home screen for Android Auto consists of two main parts: (1) a list of recent apps and functions, which appear when they are determined by the app to be relevant to the user, overlaid on top of a background graphic of a mountain landscape rendered in cool tones; and (2) a bottom menu ribbon giving five options (Navigation, Phone, Home, Music, and Exit Android Auto). Users can swipe application cards to the right to dismiss them, leaving a clean, visually appealing Android Auto home screen. Within all menus, relevant submenu options can be accessed by selecting the "drawer" – an icon comprised of three parallel, horizontal lines located in the top left-hand corner. The home button is always accessible in the center of the bottom menu ribbon.

2017 Honda Ridgeline HondaLink Center Stack. The 8-inch full-color LCD touch screen in the Ridgeline displays the standard Android Auto home screen. Adjacent to the left side of the touch screen, a column of touch-sensitive buttons provides access to both the Android Auto and native systems. Four buttons interface with Android Auto: back, audio power, a nonfunctional menu button and a home button, which exits the Android Auto app and opens the HondaLink home screen. Menu navigation and access to all features presented on the center stack display must be selected via touch.

2017 Ford Mustang SYNC 3 Center Stack. The 8-inch full-color LCD touch screen in the Mustang displays the standard Android Auto home screen. There are no dials or buttons on the center stack that can be used to control content or make selections within Android Auto. Menu navigation and access to all features presented on the center stack display must be selected via touch.

2. *Google Play Music*

Android Auto's music menu can be opened by selecting the music icon, which appears as a pair of headphones, from the bottom menu ribbon. The music selection will load with the album artwork as the full-bleed background image. Along the lower third of the screen, a menu bar provides access to the queue, skip functions, play/pause and an overflow button of additional options. In versions v2.6.573463 and v2.7.573954 of Google Music, when playing music that is stored on the phone, a playlist must first be selected from the audio drawer (the submenu denoted by three horizontal lines) to feed music into the queue. The queue is organized according to the order in which songs were added to the playlist rather than alphabetically. The alphabetical sort button allows users to jump through the queue by the first letter of the song title. After this queue jump is complete, the list is then sorted alphabetically. If the user selects a song farther down the list, all songs prior to it are removed from the queue. In order to play these removed songs, users must re-load the entire playlist again. For example, if the user selected the song "Rise" by Katy Perry, all songs that start with the letters A-Q will not be listed in the queue anymore. The user has to re-load the playlist to access these songs.

2017 Honda Ridgeline HondaLink Center Stack. When the vehicle is traveling faster than about 5 mph, the system limits users from scrolling through more than three pages of information, letting them know that "browsing [is] locked for safety" when they reach the scrolling limit. Users can cycle through the native HondaLink navigation system, the native HondaLink phone menu or the Android Auto music menu displayed on the touch screen via the menu button on the steering wheel.

2017 Ford Mustang SYNC 3 Center Stack. All selections must be made via touch screen.

3. *Calling and Dialing*

The phone menu can be accessed by pressing the phone icon on the menu ribbon along the bottom of the screen. Similar to the home screen, the phone menu includes a scrollable list of contact cards and a drawer with access to voicemail, call history, the dial pad and missed calls. The call history, dial pad and missed calls lists are all fully accessible while driving. The cards on the scrollable list are organized into three subsections: (1) the most recently made call, (2) an alphabetized list of favorite contacts, and (3) an auto-populated list of recently called contacts organized in alphabetical order. Users are shown three full contact cards, while only one-half of the fourth contact card is shown. These contact cards scroll with a "chunked" rather than smooth effect. A single tap on the contact card or number type (e.g., work or cell phone) places the call, switching the screen to the in-call menu where contact details and further options can be found. The gray dial pad is accessible while driving and can be loaded via the phone functions menu. It spans the bottom half of the screen, creating long and thin horizontal buttons. The keypad tone associated with the press of each button is consistently asynchronous with the tap of the button, motivating users to slow down the pace of phone number input. If any numbers are dialed but the call is not placed upon exiting the dialing menu, they are loaded on the dial pad again upon reopening the dialing menu. This phone number recall function applies to the last number dialed and called, meaning the user has to erase the last phone number called if they want to enter a new phone number.

2017 Honda Ridgeline HondaLink Center Stack. The designated end-call button on the steering wheel can be used to end phone calls.

2017 Ford Mustang SYNC 3 Center Stack. The phone button on the steering wheel can be used to end phone calls.

4. *Text Messaging*

Android Auto does not have a text messaging menu that is readily accessible on the touch screen. Instead, if a message is received while Android Auto is in use, an auditory notification is given and the new message appears both as a card on the home screen and as a dropdown box with sender details. If the dropdown box is selected on the touch screen, Google Assistant will read out the sender name and message contents, and then ask if users want to reply using free dictation. If users choose this option, the process to send a message is the same as in the voice interaction mode. Alternatively, the SMS card on the home screen gives the option to hear the message read aloud or reply automatically with the press of a button, "I'm driving right now." After sending this auto-reply, users are presented with the option to mute incoming message notifications.

5. ***Navigation***
The navigation menu can be accessed by pressing the arrow icon located at the far left corner of the bottom menu ribbon. This leads to a screen displaying a map centered on the vehicle's current location. Several items are overlaid onto the map, including one to two recent locations, a button to activate Google Assistant and controls to manipulate the map. To designate destinations using the touch screen, users can open the navigation drawer in the top left corner. This submenu gives users the option of selecting navigation categories or recent locations and displaying a traffic density overlay onto the map. There are four navigation categories: gas stations, restaurants, grocery stores and coffee shops. Selecting any one of these categories displays a list of 10 nearby locations within that category, organized by proximity. After users select the desired destination, an overview of the route is displayed on the touch screen. A submenu overlaid onto the route display gives users the option to either call the destination or begin the route. If users begin route guidance, the system presents auditory turn-by-turn directions and visual on-screen route guidance. Users can end route guidance by pressing the X button in the lower left corner.

Voice Comparisons

6. ***Voice Command System and Interaction***
The voice-activated Google Assistant is always available for activation via the verbal command "OK Google." Alternatively, users can press the voice command button on the steering wheel or press the microphone button in the top left corner of the touch screen to activate Google Assistant. Once Google Assistant is active, users can then give flexible, conversational commands to the voice system to select music, place phone calls, read and send text messages, and set navigation. Visual cues on the touch screen and aural cues in the form of distinct beeps inform users whether Google Assistant is speaking, listening or processing. Google Assistant and Google voice typing can be "trained" to users' voices to promote more accurate voice recognition. If users say nothing for a few seconds, Google Assistant shows a dialogue box titled, "What can you do?" This feature allows users to learn about the voice system's capabilities. After approximately nine seconds, without a command from users, the system sounds two low beeps and cancels the interaction.

2017 Honda Ridgeline HondaLink Voice Command System. A long press of the voice recognition button on the steering wheel will activate Google Assistant. Notably, a short button press will activate the native system's voice command system.

2017 Ford Mustang SYNC 3 Voice Command System. A short or long press of the voice recognition button on the steering wheel can be used to activate Google Assistant.

7. ***Google Play Music***
Google's default music app is Google Play Music, although other music apps (e.g., Pandora and Spotify) can be controlled by voice when specified. With Google Assistant, intuitive, conversational voice commands can be used to make audio selections. Users can give commands such as "Play <artist>", "I want to listen to <song>", "Let's hear some <genre>", "Play the album <album>", and many others. Google Assistant will confirm user input vocalizing the music selection that will be played (e.g. "Asking to play 'Sorry,' by Justin Bieber" or "Playing artist Louis Armstrong"). Occasionally, Google Assistant will confirm the music selection but fail to load said music. When attempting to play music stored on the phone, Google Play often begins

streaming online playlists related to user input instead. For example, if the user says, "Play 'I'm Gonna Be (500 Miles)'", the system may create a radio station based on the artist The Proclaimers but does not actually play the requested song.

8. *Calling and Dialing*

Google Assistant accepts a variety of conversational commands for placing calls to contacts and dialing phone numbers. If users have not provided enough information specifying which contact to call (e.g., there may be multiple contacts with the same name), the system asks for further specification. For instance, when Google Assistant is unable to understand a contact name, after a few attempts, the system will suggest saying only the first or last name of the contact. The system vocalizes user input as confirmation before placing a call. If Google Assistant incorrectly interprets input, users must hang up and try again.

2017 Honda Ridgeline HondaLink Voice Command System. Phone calls can be ended using the designated end-call button on the steering wheel.

2017 Ford Mustang SYNC 3 Voice Command System. Users can end phone calls by pressing the designated end-call button on the steering wheel.

9. *Text Messaging*

Google Assistant allows users to dictate new messages to contacts and phone numbers, as well as reply to unopened messages, with great accuracy. Users receive guidance from Google Assistant through the process of sending a message with succinct step-by-step instructions, a feature novice users may find particularly useful. The system is capable of parsing a string of commands, allowing more experienced users to give all message-relevant information in one command (e.g., "Text John that I'm running late"). After receiving recipient information and message content and vocalizing its interpretation, Google Assistant gives the user the opportunity to confirm or change the recipient and/or message by asking if they want to "send it, or change it?" If the system is unable to interpret the desired contact's name after multiple attempts, Google Assistant may suggest users give only the first or last name of the contact.

10. *Navigation*

Similar to its other functions, Google Assistant's voice commands for navigation are conversational and diverse. While selecting destinations from the center stack is limited to only four categories, users have more freedom when selecting destinations through the voice interface. To begin route guidance, users can give commands such as, "Navigate to <destination>", "Take me to the nearest <destination>", "Start guidance to the closest <destination>", and many others. One important note is that the addition of the keywords "nearest" or "closest" indicate for Google Assistant to automatically choose the closest location for navigation. If those keywords are not used, users are presented with a list of nearby locations as the voice interaction ends, and a selection must be made manually via the touch screen. Once a destination has been selected, users receive auditory turn-by-turn directions as well as visual on-screen route guidance. Users can cancel route guidance by activating the voice system and saying "end navigation," "cancel route guidance" or a similar command.

B. ANDROID AUTO – 2018 Ram 1500 Laramie and 2018 Chevrolet Silverado LT

Tested with the following software:
- Android Phone Operating System (OS) 8.0.0
- Android Auto App Version v2.7.573954

Center Stack Comparisons

1. *Center Stack Structure*
The home screen for Android Auto v2.7.573954 is visually and functionally the same as v2.6.573463 (refer to Appendix 2, Section A.1).

2018 Ram 1500 Uconnect Center Stack. The 8.4-inch full-color LCD touch screen in the Ram 1500 is composed of the standard Android Auto v2.7.573954 home screen in between the ever-present current status display of the native Uconnect system and the menu ribbon along the bottom of the screen. Users can navigate through certain menus and make selections via the tuning dial and enter button located below the touch screen. However, the tuning dial cannot be used to select any items from the bottom menu ribbon (e.g., the navigation, phone, music or home button). The back button in the center stack is nonfunctional while Android Auto is active.

2018 Chevrolet Silverado MyLink Center Stack. The 8-inch full-color LCD touch screen in the Silverado is composed of the standard Android Auto v2.7.573954 home screen. Users can scroll through lists and make selections via the tuning dial and menu button. This tuning dial cannot be used to select any items from the bottom menu ribbon (e.g. the navigation, phone, music or home button). The back button present on the center stack is functional within Android Auto. The designated home button loads the native system's home screen when pressed.

2. *Google Play Music*
2018 Ram 1500 Uconnect Center Stack. In addition to utilizing the touch screen directly, users can scroll through certain audio menus and the queue of music with the tuning knob and selector button.

2018 Chevrolet Silverado MyLink Center Stack. In addition to utilizing the touch screen directly, users can also scroll through certain audio menus and the queue of music with the tuning knob, selector button and back button.

3. *Calling and Dialing*
2018 Ram 1500 Uconnect Center Stack: The single steering wheel phone button cannot be used to end phone calls.

2018 Chevrolet Silverado MyLink Center Stack: The end-call button on the steering wheel can be used to end phone calls.

4. *Navigation*
2018 Ram 1500 Uconnect Center Stack: Throughout the navigation menu, the tuning dial and enter button can both be used to move through the menu structure.

2018 Chevrolet Silverado MyLink Center Stack: In addition to utilizing the touch screen directly, users can also navigate through certain navigation menus with the tuning knob, menu button and back button.

Voice Comparisons

5. *Voice Command System and Interaction*
Google Assistant for Android Auto v2.7.573954 is aurally and functionally the same as v2.6.573463 (refer to Appendix 2, Section A.6).

2018 Ram 1500 Uconnect Voice Command System. Google Assistant can be activated in the Ram 1500 by pressing and holding the voice command button on the steering wheel or by saying "OK Google." A short press will activate the native voice command system.

2018 Chevrolet Silverado MyLink Voice Command System. Users' Google Assistant can be activated by saying "OK Google," or by pressing and holding the voice command button on the steering wheel. A short press will activate the native system's voice command system.

6. *Calling and Dialing*
2018 Ram 1500 Uconnect Voice Command System. The single steering wheel phone button cannot be used to end phone calls.

2018 Chevrolet Silverado MyLink Voice Command System. The end-call button on the steering wheel can be used to cancel Google Assistant and end phone calls. (*Note*: There is only one phone-related button on the steering wheel. The check mark button is used to accept phone calls.)

C. ANDROID AUTO – 2018 Kia Optima LX

Tested with the following software:

- Android Phone Operating System (OS) 8.0.0
- Android Auto Version v2.8.5754514

Center Stack Comparisons

1. *Center Stack Structure*

The home screen for Android Auto v2.8.5754514 is functionally the same as v2.6.573463 and v.2.7.573954 (refer to Appendix 2, Section A.1). This version utilizes Android Auto's characteristic structure, with its two main parts being recent apps and functions that appear as they become relevant and the bottom menu ribbon for access to Navigation, Phone, Home, Music and Exit Android Auto menus. However, in this newer version of the app, the background design of the Home, Audio, and Phone menus has been re-designed with a higher contrast between important information and the background image, resulting in a more visually appealing and readable interface. The image of the mountain landscape featured in previous versions of the app has been replaced with a cool-toned gradient of blue and purple, presenting a clean, minimalist Android Auto home screen.

2018 Kia Optima UVO Center Stack. The 7-inch full-color LCD touch screen in the Optima is composed of the standard Android Auto v2.8.5754514 home screen. Menu navigation and all features presented on the center stack display must be selected via touch; none of the available dials and buttons can be used to control content or make selections within Android Auto. If users attempt to do so, a pop-up informational dialogue box informs the users the feature is not available while Android Auto is activated.

2. *Google Play Music*

2018 Kia Optima UVO Center Stack. The audio menu has been re-designed in this newer version of Android Auto (v2.8.5754514). Along the lower third of the screen, a menu bar provides access to the queue, skip functions, play/pause and other options. Similar to previous versions, a playlist must first be selected from the audio drawer to feed music stored on the phone into the queue. The queue remains organized according to the order in which songs were added to the playlist rather than alphabetically. The alphabetical sort button allows users to jump through the playlist or queue. The alphabetical sort button allows users to jump through the queue by the first letter of the song title. After this queue jump is complete, the list is then sorted alphabetically. New to this version of Android Auto (v2.8.5754514), prior songs in the queue no longer disappear when users select songs further down in the playlist; therefore, users are not required to re-load the playlist in order to play them (see Appendix 2, Section A.2).

3. *Calling and Dialing*

2018 Kia Optima UVO Center Stack. This version of Android Auto (v2.8.5754514) includes a re-designed phone menu, which can be accessed by pressing the phone icon on the menu ribbon along the bottom of the screen (see Appendix 2, Section A.3). The dial pad, call history and missed calls lists are accessible via the phone drawer and remain accessible while the vehicle is in motion. The familiar, scrollable list of contact cards is organized in the same manner as the previous version, though they are now clustered on the screen in groups of three, rather than three and a half. Despite the change, the phone menu is characterized by the same chunked scrolling effect present in previous versions. A single tap on the contact card or number type places the call, switching the screen to the in-call menu, which displays contact details and options. The dial pad is accessible while driving and can be loaded via the phone functions menu or a red/pink shortcut button on the main phone menu. The formerly gray dial pad is now white and is positioned on the left side of the touch screen, closest to the driver's seat. These changes increased the surface area of each button and improved its usability. Consistent with the previous version, the keypad tone associated with the press of each button is asynchronous with the tap of the button, motivating users to slow down the pace of phone number input. Upon exiting the dialing menu, all numbers remaining on the dial pad are now deleted and are no longer loaded upon reopening the dialing menu. The red end-call button can be used to hang up phone calls.

4. *Navigation*

2018 Kia Optima UVO Center Stack. All selections must be made via direct touch to the center stack touch screen, rather than using the rotary dial present in the Ram and Silverado (see Appendix 2, Section B.4).

Voice Comparisons

5. *Voice Command System and Interaction*

Google Assistant for Android Auto v2.8.5754514 is aurally and functionally the same as the previous versions discussed in this report (refer to Appendix 2, Section A.6).

2018 Kia Optima UVO Voice Command System. Users can activate Google Assistant by saying "OK Google," or by quickly pressing the voice recognition button on the steering wheel. The voice interaction can be canceled by pressing any other button on the steering wheel.

6. *Calling and Dialing*

2018 Kia Optima UVO Voice Command System. Users can answer and end phone calls using the easily recognizable green and red phone buttons, respectively, on the steering wheel.

D. APPLE CARPLAY – 2017 Ford Mustang GT and 2017 Honda Ridgeline RTL-E

Tested with the following software:
- iPhone Operating System (iOS) 10.3.3

Center Stack Comparisons

1. *Center Stack Structure*
Visually similar to the characteristic iPhone menu aesthetic comprised of color-coded square tiles, Apple CarPlay's main menu organizes popular and driving-related Apple apps (defaults include: Phone, Music, Maps, Messages, Now Playing, [return to] Uconnect, Podcasts and Audiobooks) in a grid layout against a black background for high contrast. A vertical bar on the driver side of the screen houses several items: the home button; shortcut tiles, which access the three most recently used apps; and status information, such as a digital clock and phone battery life. Unlike an iPhone, the order of the apps on the main menu is not customizable. Within any menu, a back button is positioned in the top left-hand corner when relevant. All submenus within one menu are organized cleanly along the top of the screen, avoiding the need to scroll over for more options. Depending on the capabilities of the vehicle, selections may be made by touch and/or via a tuning dial and/or selector button.

2017 Honda Ridgeline HondaLink Center Stack. The 8-inch full-color LCD touch screen in the Ridgeline displays the standard iOS 10.3.3 Apple CarPlay home screen. Adjacent to the touch screen, a column of touch-sensitive buttons provides access to both the CarPlay and native systems. Four buttons interface with Apple CarPlay: back; audio power; a menu button, which produces an audio source menu when the audio screen is open; and a home button, which exits Apple CarPlay and opens the HondaLink home screen. Menu navigation and access to all features presented on the center stack display must be selected via touch. Additionally, users can cycle through the native HondaLink navigation system and phone menu on the touch-screen display in the center stack, as well as the CarPlay music menu, by using steering wheel controls.

2017 Ford Mustang SYNC 3 Center Stack. The 8-inch full-color LCD touch screen in the Ford Mustang displays the standard iOS 10.3.3 Apple CarPlay home screen. There are no dials or buttons on the center stack that can be used to control content or make selections within Apple CarPlay. Menu navigation and all features presented on the center stack display must be selected via touch.

2. *Apple Music*
Apple CarPlay's music menu can be opened by either selecting the Music app or the Now Playing app from the main menu. If selected, the Music app loads the most recently opened submenu, whereas the Now Playing app loads the current music selection with the album artwork as the background graphic. Buttons spaced throughout the bottom half of the screen provide access to skip, play/pause, shuffle and repeat functions. When selecting music, users are presented with three categories: Library, Playlists, and Radio. Within each category, users can narrow a search by choosing different subcategories (e.g., artists, albums, songs, etc.).

2017 Honda Ridgeline HondaLink Center Stack. Several shortcuts are available in the Honda Ridgeline. While the vehicle is in motion, only the 10 most recently played items are available. If users scroll to the end of a subcategory, the system advises them to "ask Siri for more while driving." Rather than navigating through multiple submenus to find a song, users can choose to play a subcategory on shuffle.

2017 Ford Mustang SYNC 3 Center Stack. The Ford Mustang provides the same shortcut functionalities noted above for the Honda Ridgeline.

3. *Calling and Dialing*

The phone menu can be accessed by pressing the Phone app in the main menu. Immediately after selecting the app, Siri activates and asks users who they would like to call. This voice prompt can be canceled by selecting "show contacts." Users can also bypass Siri entirely by pressing the phone shortcut button in the homepage sidebar rather than the Phone app in the main menu. Five phone categories are listed along the top of the screen: Favorites, Recents [Calls], Contacts, Keypad, and Voicemail. The Favorites list is limited to 10 phone numbers that must be programmed on the phone. Names on this list require only a single tap on the contact name to place the call. The keypad consists of large circular buttons arranged to the left side of the screen near the driver. Users can place a call to the desired number by pressing the large green phone button. The contacts list is organized alphabetically by first name. Users can quickly jump to a section of the list by using the A-Z jump feature, which utilizes a full-screen alphabet. Selecting a contact produces a new screen with a list of all numbers associated with the selected contact, as well as a button to send text messages to the associated number. While engaged in a phone call, the red end-call button is positioned near the driver within easy reach.

2017 Honda Ridgeline HondaLink Center Stack. The keypad and contacts list are disabled from use while driving, which is clearly stated on the touch screen when users attempt to engage these functions.

2017 Ford Mustang SYNC 3 Center Stack. The Mustang disables the same functions as the Ridgeline as noted above when the vehicle is not in park.

4. *Text Messaging*

Text messaging is accessible on the touch screen either directly through the messages menu or through the contacts list. (Note: A system's contacts list may not be available while the vehicle is in motion — a function that varies based on decisions made by OEMs when designing the platform's ability to host Apple CarPlay.) When the messages menu is selected, the system encourages users to switch to the voice system: Siri asks users "Who do you want to text?" This voice prompt can be canceled by selecting "show messages" or bypassed entirely by using the messages shortcut button in the sidebar. The message inbox consists of a list of contact names for the 10 most recent conversations in the inbox. From here, users can select a message recipient. Alternatively, users can use the contacts list to send a text message by pressing the message icon after selecting a contact. After a message recipient is selected, Siri prompts message dictation. Once users have finished speaking, Siri repeats the dictation and asks if it should be sent. Users can select whether Siri should send, change or repeat the message via the touch screen. Commands can also be given verbally. If users opt to send the message, Siri verbally confirms this choice by saying “message sent” or giving a similar response.

2017 Honda Ridgeline HondaLink Center Stack. The contacts list is not accessible while the vehicle is in motion, requiring drivers to use the designated text messaging app to send new text messages.

2017 Ford Mustang SYNC 3 Center Stack. Like the Ridgeline, the Mustang disables the contacts list while the vehicle is in motion.

5. *Navigation*

The navigation menu can be opened by selecting the Maps app on the home screen. This leads to a screen displaying a map centered around the vehicle’s current location. Several items are overlaid onto the map, including a shortcut to navigate back to the home screen and controls to manipulate the map. To set destinations using the touch screen, users can press the Destinations button in the top right corner, loading a menu with six different navigation categories represented by familiar, intuitive icons: Recent Locations, Gas Stations, Parking, Restaurants, Coffee Shops and Grocery Stores. Selecting any of these categories displays a list of 10 nearby locations within that category organized by proximity. Additional information is listed below each destination, including street address, users’ estimated travel time, star rating, price range and distance from the vehicle. Selecting the desired destination loads a screen with an overview of the route, overlaid with more information about the destination and the option to begin route guidance. If users begin route guidance, the system provides auditory turn-by-turn directions and visual on-screen route guidance with map view and navigation options. Users can end route guidance by pressing the end button in the upper right-hand corner.

Voice Comparisons

6. *Voice Command System and Interaction*

Apple’s voice-activated assistant, Siri, is always available for activation via the verbal command “Hey Siri” through the connected device in iOS 10.3.3. Alternatively, users can press and hold the voice recognition button on the steering wheel for the length of a command to ensure Siri is listening. The voice interaction can be canceled by pressing the voice recognition button again or saying “cancel.” Users can immediately give a command after voicing, “Hey Siri,” or wait for Siri to ask for a command. Once Siri is active, the message, “What can I help you with?” is displayed on the screen. Users can give multiple conversational commands to select music, place phone calls, read and send text messages, and choose navigation. Visual cues on the touch screen and auditory cues in the form of distinct beeps inform users whether Siri is speaking, listening or processing. Siri and voice texting can be trained to users’ voices to improve accuracy. If users do not give a command for several seconds, Siri emits a chime sound, indicating it is no longer listening and has disengaged. The system is flexible and will accept an assortment of commands and formats.

2017 Honda Ridgeline HondaLink Voice Command System. Siri can be activated by pressing and holding the voice recognition button on the steering wheel – a short press activates the native system’s voice command system. The end call/back button can be used to cancel the interaction. When music is playing from a CarPlay source, users can skip through the track list, adjust volume and change the audio source to one outside of the CarPlay app via steering wheel buttons. Additionally, users can cycle through the native HondaLink navigation system and phone menu, as well as the CarPlay music menu, using these controls.

2017 Ford Mustang SYNC 3 Voice Command System. A short or long press of the voice recognition button on the steering wheel in the Mustang can be used to activate Siri.

7. *Apple Music*

Apple CarPlay's default music app is Apple Music, although other music apps (e.g., Pandora and Spotify) can be controlled via voice commands when specified. Conversational and intuitive voice commands can be used to make audio selections using Siri. Users can give commands such as "Play <artist>", "I want to listen to <song>", "Let's hear some <genre>", "Play the album <album>", and many others. Siri confirms user input by vocalizing the music selection prior to playing it (e.g. "Jazz music, coming right up" or "Here's some music by Louis Armstrong"). The Now Playing music screen is not automatically loaded after a media selection is made using Siri. Users must manually load the Now Playing screen to see media details by selecting the Now Playing app or by navigating through the Music app.

8. *Calling and Dialing*

Siri accepts a variety of conversational commands to place calls to contacts and when dialing phone numbers. If users have not provided enough information specifying which contact to call (e.g., there may be multiple contacts with the same name), the system asks for further specification. However, users need only specify the first or last name of a contact to place a call. The system vocalizes user input as confirmation before placing a call. If Siri incorrectly interprets input, the user must hang up using the red end-call button on the touch screen and try again.

2017 Honda Ridgeline HondaLink Voice Command System. Users can end phone calls using the designated end-call button on the steering wheel.

2017 Ford Mustang SYNC 3 Voice Command System. Users can end phone calls using the designated end-call button on the steering wheel.

9. *Text Messaging*

Apple CarPlay allows users to dictate new messages to contacts and phone numbers and reply to unopened messages via Siri with great accuracy. Users are guided through the process of sending a message using Siri with succinct, step-by-step instructions. The CarPlay system is capable of parsing a string of commands, allowing users to give all message-relevant information in one command (e.g., "Hey Siri, text John I'm running late"). After receiving recipient information and message contents, Siri asks the user if they want to send the message. However, the options to repeat or change the message are available only as touch-screen buttons.

10. *Navigation*

Similar to its other functions, Siri accepts conversational and intuitive voice commands for navigation. While destination selection from the center stack is limited only to six categories, users have more freedom when selecting destinations through the voice interface. To begin route guidance, users can give commands such as, "Navigate to <destination>", "Take me to the nearest

<destination>", "Start guidance to the closest <destination>", and many others. One important note is that the addition of the keywords "nearest" or "closest" indicate for Siri to automatically choose the closest location for navigation. If those keywords are not used, users are presented with a list of nearby locations as the voice interaction ends, and a selection must be made manually on the touch screen or by activating Siri again. Once a destination has been selected, users receive auditory turn-by-turn directions as well as visual on-screen route guidance. Users can cancel route guidance by activating the voice system and saying "end navigation," "cancel route guidance" or similar commands.

E. APPLE CARPLAY -- 2018 Ram 1500 Laramie, 2018 Chevrolet Silverado LT and 2018 Kia Optima LX

Tested with the following software:

- iPhone Operating System (iOS) 11.0.3

Center Stack Comparisons

1. *Center Stack Structure*
The home screen for Apple CarPlay with iOS 11.0.3 is functionally and visually the same as the previous iOS 10.3.3 tested (refer to Appendix 2, Section D.1).

2018 Ram 1500 Uconnect Center Stack. The 8.4-inch full-color LCD touch screen in the Ram 1500 displays the standard iOS 11.0.3 Apple CarPlay home screen in between the ever-present current-status display of the native Uconnect system at the top of the screen and the menu ribbon along the bottom of the screen. Users can control Apple CarPlay on the touch screen via the tuning knob and enter button located below the touch screen. The back button is nonfunctional in Apple CarPlay. In contrast to Android Auto, the tuning dial can be used to complete an entire task; it can select any item on the touch screen.

2018 Chevrolet Silverado MyLink Center Stack. The 8-inch full-color LCD touch screen in the Silverado displays the standard iOS 11.0.3 Apple CarPlay home screen. A tuning knob, menu button and back button on the center stack can be used to control Apple CarPlay on the touch screen. In contrast to Android Auto, this tuning dial can be used to complete an entire task, as it can select any item shown on the screen. Within all menus, relevant submenu options are listed across the top of the screen as static buttons and do not require scrolling. When applicable, a back button appears in the top left-hand corner. All selections can be made via touch screen or manual buttons, or a combination of the two modes of interaction.

2018 Kia Optima UVO Center Stack. The 7-inch full-color LCD touch screen in the Optima displays the standard iOS 11.0.3 Apple CarPlay home screen. There are no dials or buttons on the center stack that can be used to control content or make selections within Apple CarPlay. If users attempt to do so, an informational dialogue box informs them the feature is not available while CarPlay is activated.

2. *Apple Music*
2018 Ram 1500 Uconnect Center Stack. Several shortcuts are available in the Ram 1500. While the vehicle is in motion, only the 10 most recently played items are available within each subcategory. If users scroll to the end of a subcategory, the system advises them to “ask Siri for more while driving.” As an added shortcut while driving, users can select a subcategory — for example, a genre such as jazz — to be shuffled across all media in the category instead of navigating through multiple submenus to find a specific song. The center stack dials and buttons can be used to navigate through submenus, scroll through categories and make selections.

2018 Chevrolet Silverado MyLink Center Stack. The Silverado provides the same shortcut as the Ram 1500: Users can select a subcategory to be shuffled instead of navigating through multiple submenus to find a song. The center stack tuning knob, menu button and back button can be used to navigate through submenus, scroll through categories and make selections noted above for the Ram 1500.

2018 Kia Optima UVO Center Stack. Users have full access to the music library while the vehicle is in motion. Users are required to select a specific song (as opposed to a general category to play on shuffle; e.g., jazz music). All selections must be made via the touch screen.

3. *Calling and Dialing*

2018 Ram 1500 Uconnect Center Stack. While the vehicle is in motion, the keypad and contacts list are disabled, which is clearly stated on the touch screen when users attempt to engage these functions.

2018 Chevrolet Silverado MyLink Center Stack. The Silverado disables the same functions as the Ram 1500 when the vehicle is not parked, as noted above. The center stack tuning dial, menu button and back button can be used to navigate through submenus, scroll through categories and make selections. The phone button on the steering wheel can be used to end calls.

2018 Kia Optima UVO Center Stack. While the vehicle is not in park, the contacts list has full functionality and is available to the driver; only the keypad is disabled, which is clearly stated on the touch screen when users attempt to engage these functions. All selections must be made via the touch screen.

4. *Text Messaging*

2018 Ram 1500 Uconnect Center Stack. The contacts list is inaccessible while the vehicle is in motion, requiring drivers to select a contact from the inbox to send text messages. The center stack tuning dial and enter button can be used to navigate through submenus, scroll through categories and make selections.

2018 Chevrolet Silverado MyLink Center Stack. The Silverado requires users to use the same process for sending text messages as the Ram 1500 as noted above, with the exception of a back button available in addition to the tuning dial and menu button. The back button can be used to navigate through submenus, scroll through categories and make selections in the Silverado but not the Ram 1500.

2018 Kia Optima UVO Center Stack. The contacts list is not locked out while the vehicle is in motion; users can send new messages to any contact in the contacts list, in addition to responding to text messages in the inbox. All selections must be made via touch.

5. *Navigation*

In the updated version of iOS 11.0.3, one notable design update includes the shifting of destination information during route guidance from a white box in the lower left-hand corner to a larger black box on the left side of the screen. The black box provides higher contrast of the text against the light background of the map and additional space for larger text.

2018 Ram 1500 Uconnect Center Stack. The tuning dial and enter button can be used to scroll through lists and set the destination.

2018 Chevrolet Silverado MyLink Center Stack. The center stack tuning knob, menu button and back button can be used to navigate through submenus, scroll through categories and make selections.

2018 Kia Optima UVO Center Stack. All selections must be made via the touch screen.

Voice Comparisons

6. *Voice Command System and Interaction*
Siri for Apple CarPlay on iOS 11.0.3 is aurally and functionally the same as the previous version (iOS 10.3.3) discussed in this report (refer to Appendix 2, Section D.6).

2018 Ram 1500 Uconnect Voice System. Siri can be activated by saying “Hey Siri,” or by a long press to the voice recognition button on the steering wheel. A short press activates the native system’s voice command system.

2018 Chevrolet Silverado MyLink Voice System. Apple’s Siri can be activated by saying “Hey Siri,” or by a long press of the voice recognition button on the steering wheel. A short press activates the native system’s voice command system.

2018 Kia Optima UVO Voice System. Apple’s Siri can be activated by saying “Hey Siri,” or by a short press of the voice recognition button on the steering wheel. A long press will also activate Siri but is not necessary. The voice interaction can be canceled by pressing any other button on the steering wheel.

7. *Apple Music*
2018 Kia Optima UVO Voice System. Since the Kia Optima allows full access to the music library, there is not an option at the bottom of each music category list to activate Siri and ask for more music.

8. *Calling and Dialing*
2018 Ram 1500 Uconnect Voice System. The phone button on the steering wheel is disabled while Apple CarPlay is activated, requiring users to end phone calls using the touch screen.

2018 Chevrolet Silverado MyLink Voice System. Users can end phone calls using the designated end-call button on the steering wheel.

2018 Kia Optima UVO Voice System. Users can end phone calls using the designated end-call button on the steering wheel.

FIELD NOTES ON GLITCHES AND DESIGN CONCERNS

Researchers noted common glitches, as well as design concerns, during the testing of each version of Android Auto and Apple CarPlay. Vehicles tested with the same phone operating system and application version (if applicable) are grouped together in Sections A-E below.

A. ANDROID AUTO - 2017 Ford Mustang GT and 2017 Honda Ridgeline RTL-E

Tested with the following software:
- Android Phone Operating System (OS) 7.1.2
- Android Auto Version v2.6.573463

Glitches

1. *Google Play Execution Failure*
(1) Google Assistant: Google Play occasionally fails to load users' music selections given via voice command even after the system verbally confirms the selection. This typically occurs when the connected phone lacks adequate internet service. In the Mustang, this glitch results in an error message displayed on-screen ("Google Play is not working right now"). In the Ridgeline, Google Assistant confirms with "OK, asking to play music," and becomes stuck on the audio loading screen. After an excessive loading time (28 seconds in one instance), Google Assistant loads an incorrect music selection. In both vehicles, this type of glitch occurs whether the music selection is local to the phone (i.e. available for offline listening) or requires online streaming.

(2) Center Stack: Google Play may also fail to load the correct music selection made via touch screen in the Mustang. For example, when the song "Not Afraid" was selected by the participant, the song "Smells Like Teen Spirit" loaded instead. This error occurred three more times before the correct song played.

2. *Command Confirmation & Incorrect Execution*
Frequently, Google Assistant acknowledges users' commands but executes a different, related command. For example, when participants asked to play the downloaded song "Don't Stop Believin'" (by Journey), the system confirmed with "OK, asking to play music," and then streamed a Journey album instead.

3. *Lag Time for Verbal Commands*
Android Auto may require several minutes to complete a verbal command from users. This occurs when users give music, navigation or texting voice commands through Google Assistant.

4. *System Reset*
On at least one occasion, both the HondaLink and SYNC 3 systems crashed and automatically rebooted. In the Ridgeline, this occurred while the home screen on the Android Auto app was open and there was no task in progress. The entire episode took about 35 seconds between the time the system crashed and when Android Auto was able to be accessed. In the Mustang, the participant was in the middle of placing a call using the touch screen when the screen blacked out. The participant had to pull over and turn the vehicle off and on again to reset the system.

5. *Muted Auditory Feedback*
In the Mustang, Google Assistant sporadically muted auditory feedback to users' commands, leaving only visual feedback from the system. Occasionally, Google Assistant began providing auditory feedback when given another command. At other times, the only way users found to remedy this issue was to re-start the vehicle.

Design Concerns

6. ***Contact Card Design***
The design of contact cards in the Android Auto system is inefficient and may lead to error, due to difficulties in contact card presentation and scrolling, position on the screen, and viewing options.

(1) Users can only scroll through three and a half contact cards at a time in a chunked scrolling effect. Three full contact cards, as well as the top half of the next contact card, are presented on-screen. The last card is split across pages and is never presented as a whole, making it difficult to read. When users scroll through contacts, the bottom half of that contact card is shifted to the top of the next page.

(2) When a contact card with more than one number is selected, the additional phone numbers drop down from the contact card. In a typical interaction, the phone numbers drop down below the edge of the screen, forcing users to scroll to see the content they were expecting to have immediately visible. In a standard drop-down list design, selection of such a list shifts the content higher on the screen, either to include all content on the screen at once or to include as much content as possible. If the latter occurs, the typical design also includes a scroll bar to make users aware of the additional information and allow them to access it. In the Android Auto app, there is little visual indication that the card has expanded off-screen or that users can scroll down to access the additional phone numbers. A small, upside-down triangle on the right side of the contact card turns right-side up when the card is expanded, which may be an inadequate indication of the change for some users.

7. ***Dial Pad Design***
(1) Android Auto's confusing dial pad design hinders its usability. To access the dial pad, users must first access the drawer in the phone menu and then select the Dial a number button. Upon selection, the dial pad replaces the contacts list. This submenu consists of an entry line, standard keypad and an upside-down triangle button. Upon first glance, the function of the triangle button is unclear. Pressing it unexpectedly returns the user to the main phone screen (i.e., contacts list), which can be confusing when users dismiss the dial pad and intend to return to the drawer.

(2) The position of the dial pad in the center of the screen requires a long reach for some participants. Compounding this reach issue, some button sizes within the phone menu appear remarkably small, possibly because of the colors used on the display. The gray and black color design does not provide sufficient contrast between buttons on the dial pad. This design creates the appearance of long, thin buttons, making number selection within the dial pad more difficult.

8. ***Google Play Design***
Android Auto's Google Play features a number of unconventional design characteristics, including atypical menu navigation, content organization and accessibility. In a typical music playlist design, playlist contents are displayed on-screen after the playlist is selected to enable users to select a specific song to play. After opening the menu and selecting a playlist in Google Play, the menu unexpectedly closes and the first song in the selected list begins playing. In order to select specific

songs within the playlist, users must use the queue feature, which displays upcoming songs and is designed for quick song selection. The queue is not organized in a logical manner that would ease song selection (e.g., alphabetical by song title or artist name); rather, it is organized according to when songs were added to the playlist. This setting is not customizable. Further exacerbating user frustration, songs located in the queue prior to the one selected are removed after selection. In order to access the removed songs, users must re-load the playlist into the queue through the drawer.

9. ***Calling and Dialing – Understanding Contact Names***
Google Assistant in Android Auto frequently fails to understand contact names with ambiguous pronunciations. For example, Google Assistant frequently failed to understand the name Milly Jung. The name Jung is pronounced like Yoong or jungle without the "-le," depending on the area of origin. After a few failed attempts, the system resorted to suggesting users try saying only the first or last name. Even after users did so, Google Assistant often misunderstood both "Milly" and "Jung" and could not complete the task.

B. ANDROID AUTO – 2018 Ram 1500 Laramie and 2018 Chevrolet Silverado LT

Tested with the following software:

- Android Phone Operating System (OS) 8.0.0
- Android Auto Version v2.7.573954

Glitches

1. *Google Play Execution Failure*
This version of Android Auto did not fix the glitch noted in the previous version (v2.6.573463; refer to Appendix 3, Section A.1), regarding the system's tendency to confirm music selections but fail to load said music and only display an error message ("Google Play is not working right now").

2. *Contact Card*
Android Auto utilizes contact cards, which expand to fill the screen when selected for a phone call. During normal operation, the contact card closes and the screen returns to the contacts list when users end a call. A common glitch in this function occurs, wherein the expanded contact card does not close and the contact list reappears behind it. When this happens, the contact card cannot be dismissed, impeding the view of the contacts list and blocking access to any other phone functions. The only way users found to remedy this issue was to unplug the phone from the vehicle, effectively re-starting the Android Auto app.

3. *Muted Auditory Feedback*
This version of Android Auto did not fix the glitch noted in the previous version (v2.6.573463; refer to Appendix 3, Section A.5), regarding the phenomenon of Google Assistant auditory feedback cutting out.

Design Concerns

4. *Recurring Concerns from Previous Version*
This version of Android Auto did not fully fix the concerns noted in the previous version (v2.6.573463; refer to Appendix 3, Sections A.6-9). With respect to the contact card design (Section A.6) in the Ram and Chevrolet, using the respective touch screens to scroll through contacts still results in a chunked scrolling effect. However, the tuning dials allow users to scroll through contact cards one at a time.

C. ANDROID AUTO – 2018 Kia Optima LX

Tested with the following software:

- Android Phone Operating System (OS) 8.0.0
- Android Auto Version v2.8.5754514

Glitches

1. *System Failure*

Complete system failure occurred on two occasions while testing the navigation function. While interacting with navigation on the touch screen, the display began to noticeably lag. The navigation menu then disappeared, leaving a black box above the app menu bar. After a few seconds, the navigation menu loaded onto the screen but it was completely unresponsive. Users were able to relaunch Android Auto only after unplugging and re-connecting the phone to the car.

2. *Text Messaging – Voice Command Parsing Error*

Voice command parsing errors frequently occur when using Google Assistant to send text messages. During normal operation, the system accepts commands to activate Google Assistant and send a message to a contact as a single phrase (e.g., “OK Google, send a text message to John Doe that says, ‘Hello’”). Google Assistant often mistakenly interprets the last name of the contact as the first word of the text message (e.g., “Here’s your text to John Doe: ‘Doe hello’”). When this error occurs, users must instruct Google Assistant to change the message and then relay the message content again in order to accurately send the desired message.

Design Concerns

3. *Contact Card Design – Improvements*

Although the drop-down viewing concerns (from v2.6.573463 and v2.7.573954; refer to Appendix 3, Section A.6) remain unaddressed in the current version of Android Auto, the presentation issue is partially resolved. Contact cards are presented three at a time in a chunked scrolling effect, rather than in groups of three and a half. As a result, the system no longer splits contact cards across different pages, a change that improves readability. However, contact cards with more than one phone number are still prone to drop down off-screen, forcing users to scroll further to view the phone numbers. In fact, if the last contact card on a page is selected, the drop-down of additional numbers is completely concealed from the user.

4. *Dial Pad Design Issues – Resolved*

The updated Android Auto system resolved several issues regarding the dial pad, including the access and appearance concerns. It did not resolve the confusing design noted previously (refer to Appendix 3, Section A.7.1), which involves accessing the dial pad through the drawer. The update includes a re-designed dial pad, now positioned on the left side of the screen, within easier reach of the driver. A bright pink shortcut button to access the dial pad was added to the phone menu in the bottom right-hand corner. Furthermore, the gray and blue color palate was replaced with blue numbers against a completely white background, improving visual contrast and making number selection on the dial pad easier. This system update now automatically clears user input when exiting the dial pad instead of saving the previous entry.

D. APPLE CARPLAY – 2017 Ford Mustang GT and 2017 Honda Ridgeline RTL-E

Tested with the following software:

- iPhone Operating System (iOS) 10.3.3

Glitches

1. ***Music Selection Error***

The Apple CarPlay system occasionally experiences a glitch during music selection. During some on-road testing sessions in which Siri could not find a requested song that was stored on the phone, the system erroneously and loudly responded to users that Wi-Fi and cellular data were unavailable (e.g., "Uh-oh! Looks like you're not on Wi-Fi and haven't enabled cellular data"), despite the fact that cellular data was available at the time. This response visibly startled several participants and required researchers to briefly pause testing.

2. ***Voice Activation Failure***

Frequently, Siri fails to recognize the activation command "Hey Siri" on the first try. When this occurs, users must give the activation command again, sometimes more than once. After one or more failed attempts, many participants chose to forgo the use of the activation command and instead opted to press the voice command button on the steering wheel for more consistent results.

3. ***Premature Termination of Interaction***

Siri has been noted on several occasions to end the voice interaction prematurely. For example, after successful Siri activation via voice command, Siri beeps to indicate she is listening. She then beeps again, canceling the interaction without listening to participants' commands, often cutting them off. Another common failure occurs after Siri listens to a command from users. Siri confirms with, "OK," and then ends the interaction without executing a command or informing users why it failed.

4. ***Consistency in Interpreting Commands***

Siri often failed to interpret commands, even given in the proper format. For example, one participant took more than four minutes attempting to change the message contents of a text using Siri. The participant used the standard command "change," which typically prompts Siri to respond with, "OK, go ahead." Instead, Siri performed the review function, which reads out the message contents and asks, "Ready to send it?" When the participant replied with "no," Siri responded with, "OK, I won't send it yet. Let me know when you're ready to continue," and ended the interaction. When the participant attempted to send a message again, Siri continued to perform the review function even when the commands "cancel" and "change" were given.

5. ***Interaction Time***

Siri may take up to several minutes to complete a verbal command from users. This has been demonstrated during voice command interactions that execute navigation, text messaging and audio functions, whether or not an internet connection is required. In rare cases, the system attempts to process a single command over the entire course of the on-road experimental condition.

Design Concerns

6. *Full Access to Music Library While Driving*

In the Honda Ridgeline, users have access to the full music library while driving. This poses a concern due to the visual (eyes-off-road) and cognitive (mental) demand potentially required to reach the desired music selection. This increase in demand may result from: (1) searching through long lists of visual cues (e.g., track list); and (2) conducting multistep interactions on the touch screen.

7. *Audio Menu Design – Searching for New Music*

Within the audio menu, there is no button that takes users back to the top level of the audio menu hierarchy (i.e., the main audio page) directly from the Now Playing screen. This increases the number of steps required to start a new search.

8. *Contact Names Comprehension*

Siri often failed to understand contact names with ambiguous pronunciations. For example, Siri frequently failed to understand the name Jung, which, depending on the area of origin, is pronounced like Yoong or Jungle without the “-le”.

E. APPLE CARPLAY – 2018 Ram 1500 Laramie, 2018 Chevrolet Silverado LT and 2018 Kia Optima LX

Tested with the following software:

- iPhone Operating System (iOS) 11.0.3

Glitches

1. *Muted Auditory Feedback*

Auditory feedback from Siri was sporadically muted in the Chevrolet Silverado and Ram 1500, leaving users with only visual feedback from the system. The only way users found to remedy this issue when it occurred was to re-start the vehicle.

2. *Navigation and Route Guidance Failures*

Users encountered destination entry and route guidance initiation glitches using both the center stack and through Siri. When searching for a destination using the center stack in the Chevrolet Silverado, occasionally the system responded with a "no results" message. Participants discovered that exiting out of the navigation menu and reopening it circumvented the failure and loaded results. Every so often, when starting navigation guidance in the Ram 1500, Siri confirmed the destination but ultimately failed to begin guidance.

Design Concerns

3. *Lag Time for Route Guidance*

Siri frequently lagged between the confirmation of the destination and the start of the route guidance. Participants in the on-road study regularly chose to press the "go" button on the touch screen rather than wait for the system to begin guidance on its own. This poses a concern for users because it makes tasks take longer than necessary and may discourage the use of the hands-free voice command system in favor of the manual touch screen on the center stack.