U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

NHTSA

DOT HS 812 800

September 2019

Detection Response Task Evaluation for Driver Distraction Measurement for Auditory-Vocal Tasks: Experiment 2

DISCLAIMER

Suggested APA Format Citation:

Ranney, T. A., Baldwin, G. H. S., Skuce, I. A., Smith, L. A., & Mazzae, E. N. (2019, September). *Detection response task evaluation for driver distraction measurement for auditory-vocal tasks: Experiment 2* (Report No. DOT HS 812 800). Washington, DC: National Highway Traffic Safety Administration.

Technical Report Documentation Page

1. Report No. DOT HS 812 800	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Detection Response Task Evaluation for Driver Distraction Measurement for Auditory-Vocal Tasks: Experiment 2		5. Report Date September 2019
		6. Performing Organization Code NHTSA/NSR-120
7. Author(s) Thomas A. Ranney, G. H. Scott Baldwin, Isabella A. Skuce, & Larry A. Smith, Transportation Research Center Inc.; Elizabeth N. Mazzae, National Highway Traffic Safety Administration		8. Performing Organization Report No.
9. Performing Organization Name and Address National Highway Traffic Safety Administration Vehicle Research and Test Center P.O. Box 37 East Liberty, OH 43319		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration 1200 New Jersey Ave SE Washington, DC 20590		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code

15. Supplementary Notes

The authors thank Ed Parmer, Adam Andrella, and Jodi Welch of Transportation Research Center, Inc. for their testing support.

16. Abstract

This research evaluated the Detection Response Task (DRT) as a measure of the attentional demands of auditory-vocal in-vehicle tasks. DRT is an ISO standardized method that requires participants to respond to simple targets that occur every 3-5 s during in-vehicle task performance. DRT variants use different targets: Remote DRT (RDRT) uses visual targets; Tactile DRT (TDRT) uses vibrating targets. A single experiment evaluated the sensitivity of the two DRT variants in two test venues (driving simulator and non-driving) using auditory-vocal tasks. Participant selection criteria from the Visual-Manual NHTSA Driver Distraction Guidelines were used to recruit 192 participants; 48 were assigned to each combination of DRT variant and test venue. Identical production vehicles were used in each venue. In the simulator, participants wore a head-mounted eye tracker and performed in-vehicle tasks while driving in a car-following scenario. In the non-driving venue, occlusion testing required participants to perform the four discrete tasks while wearing occlusion goggles, which restricted viewing intermittently to simulate driving task demands. In-vehicle tasks for both venues included three discrete auditory-vocal tasks (destination entry, phone dialing, radio tuning), one discrete visual-manual task (radio tuning), and two continuous auditory-vocal digit-recall tasks representing acceptable (1-back) and unacceptable (2-back) levels of attentional load. Testing in each venue had a second part. All participants' last procedural step involved brake response time (BRT) testing in the simulator which required participants to brake in response to both expected and unexpected lead-vehicle (LV) braking events while performing selected in-vehicle tasks.

Differences observed between test venues suggest that some in-vehicle tasks are more demanding when performed intermittently in the driving simulator than when performed continuously in the non-driving venue, thus pointing to the driving simulator as the better test venue. BRT results provided some support for a connection between DRT RT and BRT; however, the experiment did not provide sufficient control of speed and headway to allow a stronger comparison. DRT results support the conclusion that the 2-back condition represents too much attentional demand and that acceptable tasks should have a lower level of attentional demand. Differences between TSOT and TEORT indicated that occlusion is not suitable for assessing auditory-vocal tasks; however, TEORT and other glance-based metrics appear suitable for use with auditory-vocal tasks. BRT testing revealed a small effect of attentional load for unexpected LV braking events but not for expected LV braking events. Mean heart rate was sensitive to differences in attentional load.

17. Key Words	18. Distribution Statement This document is available to the public through the National Technical Information Service, www.ntis.gov.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 175	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorize

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AAM	Alliance of Automobile Manufacturers
AM	amplitude modulation; pertaining to radio tuning bands
ANOVA	analysis of variance; statistical test
A-V	auditory-vocal; type of task or interface that uses voice commands
BRT	brake response time
CF	car following; driving task
CFR	Code of Federal Regulations
DAS	data acquisition system
DRT	detection response task; tool for assessing distraction
EKG	electrocardiogram; measurement by heart rate monitor
ET	eye tracker
Fam	familiarization; pertaining to name of simulated training drive
FM	frequency modulation; pertaining to radio tuning bands
GL	guidelines
HDRT	head-mounted detection response task
HR	heart rate
HRV	heart rate variability
ICF	Informed Consent Form; document describing participant rights & responsibilities
ISO	International Organization for Standardization
LV	lead vehicle; as part of a car-following scenario
N	number; sample size
N-back	delayed response task in which participants listen to and repeat simple auditory stimuli; levels include 1-back and 2-back
NADS	National Advanced Driving Simulator
OE	original equipment
OED	object and event detection
PC	personal computer
RDRT	remote detection response task
RT	response time; measure from stimulus activation to participant response
SD	standard deviation; measure of variation in a set of data
TDRT	tactile detection response task
TEORT	total eyes-off-road time
TSOT	total shutter open time
TTC	time-to-collision
V-M	visual-manual; type of task or interface that uses visual displays and physical controls manipulation
VRTC	Vehicle Research and Test Center, NHTSA’s test laboratory

EXECUTIVE SUMMARY

Distraction occurs when drivers divert their attention from the activities necessary for safe driving to a competing activity (US-EU Bilateral ITS Technical Task Force, 2010). Competing activities are generally referred to as secondary tasks, in-vehicle tasks, or tasks. NHTSA has developed voluntary guidelines to promote safety by discouraging the introduction of excessively distracting devices in vehicles. The Visual-Manual NHTSA Driver Distraction Guidelines (in 78 Fed Reg. 81, 2013 and 79 Fed. Reg. 55530; hereafter just "NHTSA Driver Distraction Guidelines") cover tasks performed with visual-manual interfaces of original equipment in-vehicle systems. The present work addresses issues related to the feasibility of extending these Guidelines to auditory-vocal tasks performed using original equipment in-vehicle systems.

Following the results of a recently completed initiative of the International Organization for Standardization (ISO, 2014), detection response tasks (DRTs) were evaluated here as a method for determining the degree to which drivers' attention was affected by auditory-vocal secondary-task demands. DRTs involve repeated presentation of a simple target while participants perform secondary tasks either in a driving simulator or in a non-driving venue with no concurrent driving task. The speed and accuracy of participants' responses to DRT targets reflect the attentional demands of the secondary tasks. Slower and/or less accurate responses have been interpreted as evidence that tasks have higher attentional load than those associated with faster, more accurate DRT responses.

The DRT has several variants, which differ primarily in their method of target presentation. Two variants were chosen for this experiment. The Remote DRT (RDRT) uses a single LED positioned away from the participant at a location near the central field of view. The Tactile DRT (TDRT) uses an electrical vibrator (i.e., tactor) taped to the participant's shoulder. One objective of this experiment was to compare the selected DRT variants for assessing the attentional load of tasks performed using auditory-vocal interfaces in a production vehicle. The ISO work (ISO, 2014) provided empirical support for the use of DRT in a non-driving venue. Accordingly, a second objective was to determine whether the use of the DRT provided consistent results in driving and non-driving test venues. A third objective was to determine whether the visual metrics (occlusion and eye-glance measures) specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks, which are performed using voice commands but may also require drivers to obtain information from in-vehicle displays.

Beyond these primary objectives, the study also sought to establish a connection between the DRT metrics and safety. The ISO work established a connection between DRT performance and the effects of attentional load (ISO, 2014); however, statistics on the effects of attentional load on crash likelihood do not currently exist. Establishing a connection between DRT response time delays and brake response time delays in emergency situations would provide a more direct link between DRT metrics and safety and thereby extend the DRT metrics' construct validity (Messick, 1995). However, recent theoretical and experimental work suggests that the behavioral mechanisms involved in responding to DRT signals, which are instructed and expected, may differ from those involved in responding to an unanticipated emergency event such as when a lead vehicle stops unexpectedly. To explore the relation between DRT and safety, a car-following scenario was included that required brake-input responses to both expected and unexpected events in the driving simulator. Lead-vehicle braking events were expected, while emergency-braking events were unexpected. The brake response time data from these events was compared with DRT response time data.

Three additional secondary objectives were identified. First, the present results were assessed to determine whether they could help establish a proposed benchmark criterion level of acceptable attentional load for auditory-vocal secondary tasks. Second, the consistency of DRT results over repeated testing was evaluated. Third, based on the work of Mehler, Reimer, and Coughlin (2012), the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks was examined.

To summarize, the study had the following objectives.

1. Compare selected DRT variants using tasks performed with auditory-vocal interfaces
2. Determine whether the use of the DRT provided consistent results in driving simulator and non-driving test venues
3. Determine whether the visual metrics (occlusion and eye-glance measures) specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks
4. Determine whether a proposed benchmark criterion level of acceptable attentional load could be established for auditory-vocal secondary tasks
5. Assess the consistency of test results over repeated testing with multiple Guidelines groups
6. Establish a connection between DRT response time and Brake Response Time (BRT) delays in emergency scenarios
7. Assess the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks
8. Determine whether increasing driving simulator test scenario task demands would influence the distributions of glance metrics used in Distraction Guidelines testing

METHOD

The study objectives were addressed in a single experiment in which test venue (simulator or non-driving venue) was a between-groups factor. One hundred ninety-two drivers participated in the experiment; 96 participants were assigned to each venue. The simulator participants completed two blocks of testing, one related to the DRT and one related to the BRT. The non-driving participants completed two blocks of testing, one related to the DRT and one related to Occlusion testing. For DRT testing, half of the participants in each venue used the TDRT and half used the RDRT.

The simulator driving scenario involved driving in the right lane of a four-lane roadway while maintaining a specified following distance in one of two assigned lead-vehicle speed conditions. Constant car-following required drivers to maintain a 220-foot following distance while the lead vehicle maintained a constant speed. Complex car-following used a changing lead-vehicle speed and required drivers to maintain a 120-foot following distance. Non-driving participants performed in-vehicle tasks and DRT in a stationary vehicle with no driving task.

Participants were recruited following NHTSA Distraction Guidelines specifications to form eight 24-person samples, four groups in each venue. For each group of 24 participants, 6 participants (3 females and 3 males) were recruited in each of the following age ranges: 18-24, 25-39, 40-54, and 55 and older.

Contemporary in-vehicle information systems facilitate a variety of activities, many of which are considered unbounded, or incompatible with the testable task definition in NHTSA Distraction Guidelines, which requires a sequence of control operations intended to accomplish a stated objective.[1] In the DRT testing component of the experiment, all participants performed six tasks in their respective combinations of test venue and DRT. Tasks included four discrete in-vehicle tasks, three of which were performed using voice commands (destination entry, radio tuning, and phone dialing) and one performed using a visual-manual interface (visual-manual radio tuning). These tasks are consistent with the definition of a testable task. Additional tasks included two levels of the N-back task,[2] a continuous delayed digit-recall task, which used auditory stimuli and required verbal responses. Based on previous research, the level of attentional demand associated with the 1-back task is generally considered acceptable for performance while driving. The attentional demand associated with 2-back is generally considered unacceptable for performance while driving. Inclusion of these tasks was intended to provide benchmark levels of attentional demand.

Summary of Findings: Detection Response Tasks

This section summarizes findings that relate to study objectives one, two, four and five. The first study objective was to compare selected DRT variants using tasks performed with auditory-vocal interfaces. The second study objective was to determine whether the use of the DRT provided consistent results in driving simulator and non-driving test venues. The fourth study objective was to determine whether a proposed benchmark criterion level of acceptable attentional load could be established for auditory-vocal secondary tasks. The fifth study objective was to assess the consistency of test results over repeated testing with multiple Guidelines groups.

1. In the driving simulator, responses to TDRT targets were consistently slower, less accurate, and more variable than responses to RDRT targets. These differences were not apparent in the non-driving venue. DRT accuracy in the non-driving venue was constrained by the ceiling of perfect performance, which reduced this metric’s sensitivity for detecting differences among task conditions.
2. The results revealed differences between DRT variants. Planned comparisons between task and benchmark conditions (1-back and 2-back) revealed differences between DRT

---

[1] Testable Task means a sequence of control operations performed using a specific method leading to a goal toward which a driver will normally persist until the goal is reached. A testable task begins with the device at a previously defined start state and proceeds, if the testable task is successfully completed, until the device attains a previously defined end state.

[2] The N-back task is described in detail in Section 2.6.1, which includes an example in Table 2.

variants in both venues. The differences were most apparent in the simulator for the RT metric, for which 6 of 9 comparisons had different outcomes for the two DRT variants. Differences between DRT variants were less pronounced in the non-driving venue.

3. The ordering of DRT RT mean values for the discrete in-vehicle tasks differed between test venues. The RT values were considerably slower relative to the benchmark conditions in the simulator than in the non-driving venue suggesting that the tasks that involved monitoring in-vehicle information screens were much more demanding when they were performed intermittently while driving than when performed continuously in the non-driving venue.
4. The two DRT metrics (RT and accuracy) provided different outcomes in some comparisons between task conditions. Approximately half of the comparisons between task conditions yielded different statistical test outcomes for DRT RT and DRT accuracy. The ISO standard recommends using both metrics for assessing tasks; however, in practice the RT results are generally given stronger weight.
5. Age effects were present, reflecting the expected finding that response times increase and DRT accuracy decreases with increasing age. Age x Task interactions were found in both simulator metrics but not in the non-driving metrics, indicating that the ordering of task conditions differed for the different age groups in the simulator.
6. Contrary to predictions, differences among samples were apparent. Differences were observed between DRT types in the simulator. In the non-driving venue, differences between samples were less apparent.

Summary of Findings: Brake Response Time

This section summarizes the findings of the sixth study objective, which was to establish a connection between DRT response time and brake response time delays in emergency scenarios.

The BRT component of the experiment followed DRT testing in the simulator venue. In the BRT component, simulator participants performed visual-manual radio tuning, 1-back, and 2-back tasks, and a baseline trial in a car-following scenario. Each driving trial had six lead-vehicle braking events with brake-light activation and minor deceleration, followed by a final event in which the lead vehicle came to a complete stop with no brake-light activation. Participants were instructed to brake immediately in response to lead-vehicle brake-light activation, but were given no instructions about the final unexpected event. During the scenario, the lead vehicle maintained a constant speed of 50 miles per hour when not braking. Participants received auditory feedback if their following distance exceeded the specified target value.

1. BRTs to expected lead-vehicle braking events were generally faster (*M* = 1.29 seconds) than responses to unexpected lead-vehicle stopping events with no brake light activation (*M* = 1.93 seconds). When expected lead-vehicle braking events were considered in the aggregate, mean BRTs associated with the 1-back, 2-back, and baseline conditions were approximately equal, suggesting that increased attentional load did not affect the brake response time to these expected events.
2. The mean BRT to expected lead-vehicle events was elevated for visual-manual radio tuning relative to baseline. Half of the expected braking events occurred when the participant was actively tuning the radio and half occurred when the participant was not actively tuning the radio. The mean BRT during task performance was 1.54 seconds versus 1.38 seconds when not performing the task. The elevation of the latter value

relative to the baseline BRT (1.38 seconds vs.1.23 seconds) likely reflects the residual attentional demand associated with task performance that was just completed.

3. BRT means associated with unexpected events revealed small effects consistent with slower responses due to increased attentional load. Specifically, the baseline (no load) mean BRT was faster than the 2-back (high load) and 1-back (medium load) means. Responses to the first unexpected event were slower, reflecting the increased uncertainty (surprise) associated with this event. Differences between conditions for this event were also consistent with effects of attentional load, although they were not strong.
4. The experiment did not allow for precise control of speed and headway at the start of the lead-vehicle braking events. Additional analyses were conducted in which the events were separated by following distance at event onset. For the unexpected events, BRTs were influenced by headway; faster BRTs were associated with shorter headways. Among the subset of unexpected trials with headways close to the target value (mid-range), BRT values appeared to reveal an effect of attentional load with shorter BRTs associated with baseline and longer BRTs associated with 2-back. Among the subset of trials with shorter headways, there was no apparent effect of attentional demand, reflecting a pattern consistent with the predictions concerning looming effects. However, the shorter-headway trials are not all unexpected, which weakens the assumption that looming cues were the only factor influencing BRT. Visual cues associated with the decreasing distance to the lead vehicle attained the looming threshold for both subsets.

Summary of Findings: Occlusion

This section summarizes findings that relate to the occlusion portion of the third study objective, which was to determine whether the occlusion measure specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks.

Participants assigned to the non-driving venue completed the occlusion test procedure in addition to the DRT testing. The occlusion testing followed the protocol specified in the NHTSA Driver Distraction Guidelines. For each of the four in-vehicle tasks, participants completed five trials while seated in the stationary vehicle and wearing occlusion goggles. The goggles could be made to be either occluded (closed) or unoccluded (open). The protocol used alternating 1.5 seconds occluded and unoccluded intervals during task performance. The performance metric was the total amount of unoccluded time during task performance, or total-shutter-open time.

1. Contrary to prediction, TSOT values for two auditory-vocal tasks (destination entry and phone dialing) were greater than for visual-manual radio tuning, which required more off-road glances for task completion. TSOT values were more consistent with task-completion time means than results from analysis of glance data (see below).
2. Computing TSOT directly produced TSOT values that were on average 0.5 seconds less than those based on the number of shutter-open intervals. Longer TSOT values associated with the open-interval computation method derive from the inclusion of full open intervals with part of the last interval occurring after the task was completed on some trials.
3. Precisely defining task-completion time is difficult; it includes time for participants to report being done and time for the experimenter to press a button to mark the completion time. Some participants may not fully appreciate the need for a timely utterance to

indicate task completion. Similarly, vigilance among experimenters is required to minimize unwanted delay.

Summary of Findings: Glance Analyses

This section summarizes findings that relate to the eye glance portion of the third study objective and the eighth study objective. The eye glance portion of the third study objective was to determine whether the eye-glance measures specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks. The eighth study objective was to determine whether increasing the driving simulator test scenario task demands would influence the distributions of glance metrics used in the NHTSA Distraction Guidelines testing.

1. Participants assigned to the driving simulator wore head-mounted eye-tracking glasses to measure and record glance behavior while driving. The performance metric was the Total amount of Eyes-Off-Road Time (TEORT) during a task. Half the simulator participants performed a simple car-following task (constant lead-vehicle speed) and the other half performed a more demanding car-following task (complex; systematically varying lead-vehicle speed), which required participants to maintain a closer following distance while the lead vehicle speed varied systematically. Visual-manual radio tuning had a higher mean TEORT than the auditory-vocal tasks. This is consistent with the expectation that visual-manual tasks require more off-road glances than auditory-vocal tasks.
2. Mean TEORT values by task were not consistent with mean TSOT values obtained in the occlusion paradigm.
3. Although not statistically significant, constant car-following, which represents a low-demand driving situation, was associated with slightly higher mean values for TEORT, mean glance duration, and proportion of long glances (> 2.0 seconds) relative to the more demanding complex car-following situation. This pattern was most pronounced in the visual-manual radio tuning condition, which required off-road glances for task performance. The effect of increasing driving task demands was most evident in reducing the overall variance associated with the distribution of responses.

Summary of Findings: Heart Rate

This section summarizes findings that relate to the seventh study objective, which was to assess the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks.

Heart rate was collected continuously during all trials in both the simulator and non-driving venues. The performance metric was mean heart rate (number of beats per minute). Data from the N-back (2-back, 1-back) and baseline trials were analyzed.

1. Mean heart rate was sensitive to changes in cognitive demand associated with N-back in both venues. This is consistent with previous findings that heart rate increases as demand and workload increases (Reimer & Mehler, 2011; Mehler, Reimer, & Coughlin, 2012).
2. Heart rate variability revealed sensitivity to increased cognitive demand, but the effects were not consistent across venues.

Conclusions

1. Responses to visual stimuli associated with the RDRT were generally faster, more accurate, and less variable than responses to the vibrating stimuli associated with the TDRT in the simulator. Although these differences were not consistent with findings of previous related research, they suggest that the RDRT had greater sensitivity and more consistency in detecting targeted differences than the TDRT in the driving simulator. Differences between DRT variants in the non-driving venue were minimal, suggesting that either DRT could effectively be used in that venue (Objective 1).
2. The driving simulator and non-driving venues did not give consistent results for some tasks. Non-driving results appear to represent the relative difficulty among tasks performed continuously without interruption. Simulator results reflect the relative difficulty among tasks performed intermittently while driving. Simulator test results are more likely to generalize to on-road driving than non-driving test results (Objective 2).
3. DRT RT is more valuable as a metric than hit rate (accuracy), which is limited by the ceiling of perfect performance, most often in the non-driving venue. however, monitoring DRT accuracy is necessary to identify speed-accuracy tradeoffs and noticeably poor DRT performance. Participants did not trade accuracy for speed in this experiment. however, DRT metrics occasionally provided inconsistent results in comparisons between conditions (Objectives 1 and 2).
4. The N-back conditions provide the strongest foundation for defining a threshold of acceptable attentional demand for assessing auditory-vocal tasks. The conclusion that 2-back represents an unacceptable level of attentional demand supports a decision model that could require acceptable tasks to have a significantly lower level of attentional demand than that of the 2-back task (Objective 4).
5. BRT results suggest that DRT RT and BRT to unexpected lead-vehicle stopping may both be sensitive to effects of attentional load; however, the inability to precisely control speed and headway at the start of the lead-vehicle braking events in the Guidelines car-following scenario precluded a strong conclusion. Stronger controls of headway between vehicles at braking-event onset than those provided by the Guidelines test protocol will be necessary to test the relation between BRT and DRT RT and thus the relation of DRT performance to safety (Objective 6).
6. Glance metrics obtained in the driving simulator (TEORT) can effectively be used to assess visual demands of auditory-vocal tasks performed with voice commands and verbal system feedback. However, occlusion is not suitable for use with these tasks. TSOT values obtained in the occlusion paradigm were not consistent with TEORT values obtained from glance data analysis. Occlusion is only suitable for assessing tasks with constant visual demand during task performance (Objective 3).
7. The visual-manual radio tuning task, as performed in the Ford Explorer test vehicle used in this study, showed a high error rate in task performance. Specifically, 39 percent of participants (37/96) had three or more error trials. The NHTSA Driver Distraction Guidelines protocol calls for replacement of participants without a specified limiting provision. It is only after 24 participants have successfully completed the task that test users compute aggregate scores to determine whether the task should be determined to be “unreasonably difficult” based on the NHTSA Driver Distraction Guidelines criterion of 50 percent errors.

In summary, the driving simulator venue combined with RDRT had greater sensitivity and more consistency in detecting targeted differences than the other venue/DRT combinations. The results support the conclusion that 2-back represents an unacceptable level of attentional demand for tasks to be performed while driving. Lastly, results showed that increased driving scenario driving task demands allow for better control of off-road glance durations.

1.0 INTRODUCTION

Distraction occurs when drivers divert their attention from the activities necessary for safe driving to a competing activity (US-EU Bilateral ITS Technical Task Force, 2010). Competing activities are generally referred to as secondary tasks, in-vehicle tasks, or tasks. NHTSA has developed voluntary guidelines to promote safety by discouraging the introduction of excessively distracting devices in vehicles. The Visual-Manual NHTSA Driver Distraction Guidelines addressed tasks performed using original equipment in-vehicle devices with visual-manual interfaces. The present work addresses tasks performed using in-vehicle devices capable of auditory-vocal interactions. Because auditory-vocal tasks may pose different demands on drivers' attention than those covered in the NHTSA Driver Distraction Guidelines, this research examines whether different metrics are warranted for assessing auditory-vocal task conformance with the NHTSA Guidelines.

1.1 Detection Response Tasks

Following the results of an ongoing initiative of the International Organization for Standardization (ISO, 2014), DRTs are being evaluated as a method for determining the degree to which drivers' attention is affected by secondary task demands. DRTs involve repeated presentation of a simple target while participants perform secondary tasks either in a driving simulator venue or in a non-driving venue with no concurrent driving task. The speed and accuracy of participants' responses reflect the attentional demands of the secondary tasks. Slower and/or less accurate responses are interpreted as evidence that tasks have higher attentional load than those associated with faster, more accurate DRT responses.

The DRT has several variants, which differ primarily in their method of target presentation. The head-mounted DRT consists of a single LED attached to a fixture worn on the head. The Remote DRT (RDRT) uses a single LED positioned away from the participant in a location near the central field of view. The Tactile DRT (TDRT) uses an electrical vibrator (i.e., tactor) taped to the participant's shoulder. When the participant either sees the LED light up or feels the electrical vibrator (the target presentation), the participant presses a button attached to their finger to record their response to the stimulus. In a recently completed experiment, hereafter referred to as Experiment 1, these three DRT variants were evaluated in driving and non-driving test venues (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014). Secondary tasks included a continuous delayed digit-recall task (N-back) and visual-manual radio tuning. The N-back task uses auditory stimuli and requires verbal responses to simulate the demands of simple conversations with different attentional loads. The 1-back level represents a simple conversation with low demand. The 2-back level represents a more demanding conversation. The main objective of Experiment 1 was to determine which DRT variants were most sensitive in detecting targeted differences among the secondary task conditions in each test venue.

The results of Experiment 1 revealed minor differences among the DRT variants. All three DRT variants were generally successful in detecting the targeted differences. The TDRT provided slightly greater sensitivity than the other DRT variants in the driving simulator. However, there were no substantive differences among DRT variants in the non-driving venue.

Based on these results, the TDRT was recommended for additional testing in the driving simulator. The TDRT is also preferred based on theoretical and methodological reasons because it does the best job of eliminating potential conflicts that can occur when DRT variants with visual targets are used to assess tasks that have visual demands (e.g., radio tuning). Briefly, while all DRT variants provide relatively pure measures of attentional load when used to assess tasks that are entirely voice-based, the metrics associated with visual target DRT variants can also contain varying contributions due to visual conflicts when they are used to assess tasks with visual demands and the target is not immediately visible when first activated. These potential conflicts are described in greater detail elsewhere (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014). The use of a tactile target eliminates visual conflicts and allows the TDRT to isolate the effects of attentional load.

The results of Experiment 1 did not support selection of a single DRT variant for use in the non-driving venue. However, based on both theoretical and practical considerations, the RDRT was selected for additional testing in the non-driving venue. The RDRT is considered to have an advantage in terms of both ecological validity (Merat & Jamson, 2008) and face validity because the remote target presentation appears consistent with the demands of hazard detection in real-world driving. Arguments based on face validity have enhanced the credibility and facilitated the widespread use of DRT and its predecessors. Thus, the RDRT has a stronger empirical foundation than the newer DRT variants that fix the target location relative to the driver. And while the use of a remote target location can introduce conflicts between DRT and secondary task demands when the RDRT is used to assess tasks performed using visual-manual interfaces, the results of Experiment 1 suggest that the effects of this potential conflict may be too small to affect differences between secondary task conditions.

Therefore, the first objective of Experiment 2 was to compare the selected DRT variants for assessing tasks performed using auditory-vocal interfaces in a production vehicle.

1.2 DRT Consistency, Simulator and Non-Driving Test Protocols

The results of Experiment 1 revealed consistent differences between the two test venues. DRT response times were consistently faster and less variable in the non-driving venue. DRT hit rates in the driving simulator were sensitive to most targeted differences; however, in the non-driving venue consistently near-perfect target-detection performance significantly reduced the sensitivity of hit rate for assessing differences across conditions. The differences between venues were attributable to the absence of a concurrent driving task, which made target detection easier in the non-driving venue. Thus, although the driving simulator provides a more valid representation of the concurrent demands of driving and secondary task performance than the non-driving venue, the consistent overall pattern of response time differences in the two venues supported the conclusion that this metric provides adequate relative validity to replicate driving-simulator test results. Therefore, the second objective of this experiment was to determine whether the driving and non-driving test protocols provided consistent results when assessing differences between tasks performed with an auditory-vocal interface.

1.3 NHTSA Driver Distraction Guidelines Metrics

The third objective of this study was to determine whether the visual metrics of the NHTSA Driver Distraction Guidelines (i.e., occlusion and eye-glance measures) could be used to assess in-vehicle tasks performed using devices with auditory-vocal interfaces. The specific tasks assessed in Experiment 2 were defined to be consistent with the NHTSA Driver Distraction Guidelines definition of a testable task,[3] which requires a sequence of control operations intended to accomplish a stated objective. Thus, while contemporary in-vehicle information systems facilitate a variety of activities, some of these activities are considered unbounded, or incompatible with the testable task definition. For example, while contemporary interfaces allow drivers to engage in phone calls using voice commands, the focus of the current experiment is on the initial command sequence that allows the driver to initiate the phone conversation using voice commands. The actual content of the phone conversation is considered out of scope because neither the specific task objective nor the defined end state can be consistently defined.

1.4 Benchmark of Acceptable Level of Attentional Load

The NHTSA 's Visual-Manual Driver Distraction Guidelines are based on the premise that visual-manual radio tuning represents the upper limit of acceptable visual-manual demand for tasks performed while driving. Because visual-manual radio tuning has an established connection to safety (Stutts, Reinfurt, Staplin, & Rodgman, 2001; Wang, Knipling, & Goodman, 1996; Wierwille & Tijerina, 1998), and to eliminate confusion about having multiple bases underlying criterion values, this research explores the feasibility of using a representative level of visual-manual radio tuning to serve as the criterion level of attentional load for use with DRT metrics. The results of Experiment 1 provided initial support for the use of visual-manual radio tuning as a benchmark criterion level of attentional load. Specifically, DRT results indicated that radio tuning has an attentional load that is higher than that associated with 1-back (see 2.6.1), which is generally considered acceptable for tasks performed while driving. Thus, if the DRT performance associated with visual-manual radio tuning were used as a criterion, 1-back would have been considered acceptable for performance while driving. This conclusion is consistent with previous research results (Ranney, Baldwin, Parmer, Domeyer, Martin & Mazzae, 2011) and those of others (Reimer, Mehler, Dobres, & Coughlin, 2013). However, it is possible that some portion of the difference in DRT performance between 1-back and visual-manual radio tuning may be due to the overhead activities associated with repeated performance of visual-manual radio tuning that are not part of the 1-back task, which requires continuous uninterrupted performance. The fourth objective of Experiment 2, therefore, was to determine whether a proposed benchmark level of acceptable attentional load could be established for auditory-vocal secondary tasks.

---

[3] Testable Task means a sequence of control operations performed using a specific method leading to a goal toward which a driver will normally persist until the goal is reached. A testable task begins with the device at a previously defined start state and proceeds, if the testable task is successfully completed, until the device attains a previously defined end state.

1.5 Consistency of Test Results Over Repeated Testing

The results of Experiment 1 revealed the potential for inconsistent test results in repeated testing using different groups of participants (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014). Therefore, the fifth objective of this study was to assess the consistency of test results over repeated testing with multiple groups of participants selected using NHTSA Driver Distraction Guidelines age and gender criteria.

1.6 Brake Response Time in Emergency Scenarios

When a distracted driver's attention is diverted from activities critical for safe driving, situations are more likely to occur in which the driver either misses or is slow in responding to events that occur on the roadway. Numerous studies have demonstrated that distraction results in slowed response time (e.g., Horrey & Wickens, 2006; Caird, Willness, Steel, & Scialfa, 2008). Two primary methods have been used to assess this effect. One method, called object and event detection, involves recording drivers' responses to unexpected hazards, such as a stopped vehicle in the roadway ahead (e.g., Lee, McGehee, Brown, & Reyes, 2002; Strayer, Drews, & Johnson, 2003). OED methods typically present realistic scenarios to elicit the surprise inherent in drivers' responses to unexpected critical events. However, because it is not possible to surprise participants repeatedly (Engström, Aust, & Viström, 2010), the OED methods that rely on responses to unexpected events are not suitable for use in test protocols that require repeated testing to facilitate comparisons across multiple task conditions. DRT methods, which use frequent presentations of targets that are less unexpected, represent the second method for assessing the slowing of responses among distracted drivers. Although less realistic, DRT methods are more suitable for testing that requires comparisons across task conditions. DRT methods are supported by a long history of research demonstrating their sensitivity to differences in levels of attentional demand (Victor, Engström, & Harbluk, 2009). The strong empirical foundation together with the practical advantages of DRT methods led to the ISO decision to select the DRT as the most promising method for assessing differences in the potential for distraction due to differences in attentional demand between tasks (ISO, 2014).

The ISO work has established a connection between DRT performance and the effects of attentional load (ISO, 2014); however, statistics on the effects of different levels of attentional load on crash likelihood do not currently exist. NHTSA sought to establish a connection between the proposed test protocols and safety to the extent possible. Response time is an integral component of drivers' responses in many situations; delayed brake responses may increase the likelihood of a crash outcome. In DRT test protocols, the effects of increasing levels of attentional load are revealed as response time delays. Establishing a connection between DRT response time delays and Brake Response Time (BRT) delays in emergency situations would provide a more direct link between DRT metrics and safety and thereby extend the DRT metrics' construct validity (Messick, 1995). However, recent theoretical and experimental work suggests that the behavioral mechanisms involved in responding to DRT signals differ from those involved in responding to an unanticipated emergency situation such as when a lead vehicle stops unexpectedly. In particular, Engström (2014), using the distinction made by Schneider and Shiffrin (1977), has argued that responding to unanticipated emergency situations is more likely to involve automatic responses while DRT responding involves more controlled processing. This distinction is important in Engström's conceptualization because it supports findings that have shown that unlike the DRT, BRTs in unexpected situations are not affected by the driver's attentional load. If this is true, then it would be difficult to establish a direct connection between these two tasks, which would mean that DRT responding is not predictive of BRT in unexpected emergency scenarios. However, it is not entirely clear what proportion of crashes involves situations that turn critical without any advance warning. Automatic responding to lead-vehicle braking appears most resistant to effects of cognitive load when there are no secondary cues that might engage the drivers' cognitive apparatus. Indeed, results of some studies have shown that LV brake-light activation may be sufficient to alert drivers to the possibility of a collision-imminent situation. The extent to which drivers must have advance warning to engage attentional resources is not well established.

Of particular interest, there may be a relation between DRT responding and BRT in driving situations that differ from unanticipated LV braking situations. According to Engström's analysis, responding to DRT tasks is a controlled task that requires cognitive activity and is thus sensitive to differences among secondary tasks in cognitive load. In contrast, per his analysis, responding with brake input in emergency situations involves very little cognitive activity and is thus not sensitive to the effects of different cognitive loads. In this latter situation, the task of responding to an unexpected LV braking situation is native and directed by the looming cues that engage (self-protective) processes that are highly automatic.

Accordingly, the sixth objective of Experiment 2 will be to obtain data necessary to provide a link between DRT target response delay and BRT in a simulated emergency scenario.

1.7 Heart Rate as an Indicator of Attentional Task Demand

Research has shown that physiological measures can be used to detect variations in the attentional demand associated with an auditory-vocal working memory task. In a series of on-road experiments, Mehler, Reimer, and colleagues demonstrated that both heart rate and skin conductance level were sensitive to incremental differences in the attentional load associated with different levels of the N-back task (Mehler, Reimer, Coughlin, & Dusek, 2009; Mehler, Reimer, & Coughlin, 2012). The seventh objective of Experiment 2 was to explore the usefulness of heart rate as a measure of attentional demand associated with auditory-vocal tasks performed in both driving and non-driving venues.

1.8 The Effect of Driving Task Demands on Glance Behavior

Research indicates that drivers adapt the durations of their off-road glances to match the demands of the driving situation (Wierwille, 1993; Tsimhoni, Yoo, & Green, 1991). In more demanding situations, drivers use shorter off-road glances. In response to NHTSA Distraction Guidelines simulator test specifications, Kujala, Lasch, and Makela (2014) argued that the long-glance criterion should be tied more closely to the driving task demands of the simulator test scenario. They used voluntary occlusion[4] to demonstrate that 46 percent of a sample of simulator

---

[4] Voluntary occlusion is a technique that measures the amount of time drivers look away from the forward roadway. With occlusion goggles closed, drivers request a view of the roadway whenever they think it is necessary for safe vehicle control. The time between successive requests is the voluntary occlusion time. Because the distribution of voluntary occlusion times represents the distribution of times a driver feels comfortable looking away from the

participants who drove a scenario like that specified in the NHTSA Driver Distraction Guidelines had 85th percentile occlusion time values greater than 2 seconds. They argued that the percentage of test participants with 85th percentile voluntary occlusion time values greater than 2.0 seconds should be closer to 15 percent, reflecting the fact that the Distraction Guidelines long-glance criterion is derived from the 85th percentile of off-road glance durations. Based on their finding, they concluded that the driving task demands of the NHTSA Driver Distraction Guidelines simulator scenario were too low. They demonstrated that increasing the scenario driving task demands reduced the voluntary occlusion time. The resulting distributions had fewer participants (38%) with 85th percentile values greater than 2.0 seconds. Therefore, the eighth objective of the present study was to determine whether increasing the driving task demands of the simulator test scenario would influence the distributions of glance metrics used in the Distraction Guidelines testing.

1.9 Data Collection Interval

DRT testing requires a constant level of task demand over the designated data collection interval. Results of Experiment 1 identified 2 minutes as the minimum data collection interval required for optimal metric sensitivity. However, one effect of performing tasks using voice commands is that the time required for a single task instance will be reduced considerably relative to its visual-manual counterpart. For example, while visual-manual radio tuning was found to require 20 to 30 seconds, it was anticipated that the voice-based version of this task required approximately 5 s. When such tasks are performed continuously over a 2-min data collection interval, a significant portion of the interval will be devoted to non-task-related or ‘overhead’ activities, which include processing task instructions and transitioning between instructions and task initiation. For extremely short tasks, such as radio tuning, the overhead activities may consume 50 percent or more of the data collection interval. The effect of these overhead activities on DRT performance is generally unknown, but it is possible that the attentional demands associated with transitions between task-related and overhead activities could be greater than those associated with the tasks.

To address this potential problem, data collection intervals were structured to allow separation of overhead from task-related time segments. Task-related time segments were combined into a single integrated file that allowed computation of mean response times and hit rates over a composite 2-min interval. For extremely short tasks (i.e., with durations near 5 s), it was necessary to ensure that each task instance had at least one DRT target presentation. Task instances that are completed without a single DRT target presentation are not consistent with ISO specifications. This was accomplished by requiring the DRT control program to schedule a DRT target during the first few seconds of each task instance.

Three-minute data collection intervals were used in the experiment. New task instances were presented at fixed intervals, which were defined to allow participants to have a short (5-10 s) rest between task instances. The rest interval allowed the driver to focus exclusively on the driving task and thus reduce the possibility that the demands of secondary task instances had lingering and potentially cumulative effects.

---

roadway, it can be taken to represent the distribution of off-road glances that could comfortably be used to perform in-vehicle tasks.

1.10 Study Objectives

To summarize, Experiment 2 had the following objectives:

1. Compare selected DRT variants using tasks performed with auditory-vocal interfaces
2. Determine whether the use of the DRT provided consistent results in driving simulator and non-driving test venues
3. Determine whether the visual metrics (i.e., occlusion and eye-glance measures) specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks
4. Determine whether a proposed benchmark criterion level of acceptable attentional load could be established for auditory-vocal secondary tasks
5. Assess the consistency of test results over repeated testing with multiple Guidelines groups with auditory-vocal tasks
6. Establish a connection between DRT response time and brake response time (BRT) delays in emergency scenarios
7. Assess the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks
8. Determine whether increasing driving simulator test scenario task demands would influence the distributions of glance metrics used in Distraction Guidelines testing

2.0 METHOD

2.1 Participants

One hundred ninety-two drivers participated in the experiment. They were recruited following NHTSA Driver Distraction Guidelines test participant recommendations to form eight 24-person samples. For each group of 24 participants, 6 participants (3 females and 3 males) were recruited in each of the following age ranges: 18-24, 25-39, 40-54 and 55 and older. Participants had to report being in good general health, have an active valid driver's license with no vision or hearing problems, have driven 3,000 miles in the last year, have experience using a wireless phone while driving, and be unfamiliar with the technology being evaluated.

2.1.1 Recruiting

Participants were recruited through advertisements placed in local Ohio newspapers, including those in Marysville, Bellefontaine, Delaware, and Kenton. Ads were placed both in physical newspapers and on newspaper websites. Recruiting materials are presented in Appendix A. To facilitate recruitment, an online application procedure was used, which allowed potential participants to complete the screening questionnaire online. Information obtained in this manner was downloaded routinely for use in assessing the eligibility of respondents and for scheduling participation. A phone number was also provided in the advertisements as an alternative to the online application, in case there was anyone who was unable to respond using the online method. The participant screening questions are presented in Appendix B.

2.1.2 Informed Consent

For those selected for participation, an informed consent was obtained in accordance with 45 CFR 46.116. The informed-consent form is included in Appendix C. Participants were also asked to sign an information disclosure statement, also included in Appendix C, which gave NHTSA the right to use the engineering, video, and audio data from the study for scientific, educational, research or outreach purposes. The procedure used for obtaining informed consent is described in Section 2.7.1.

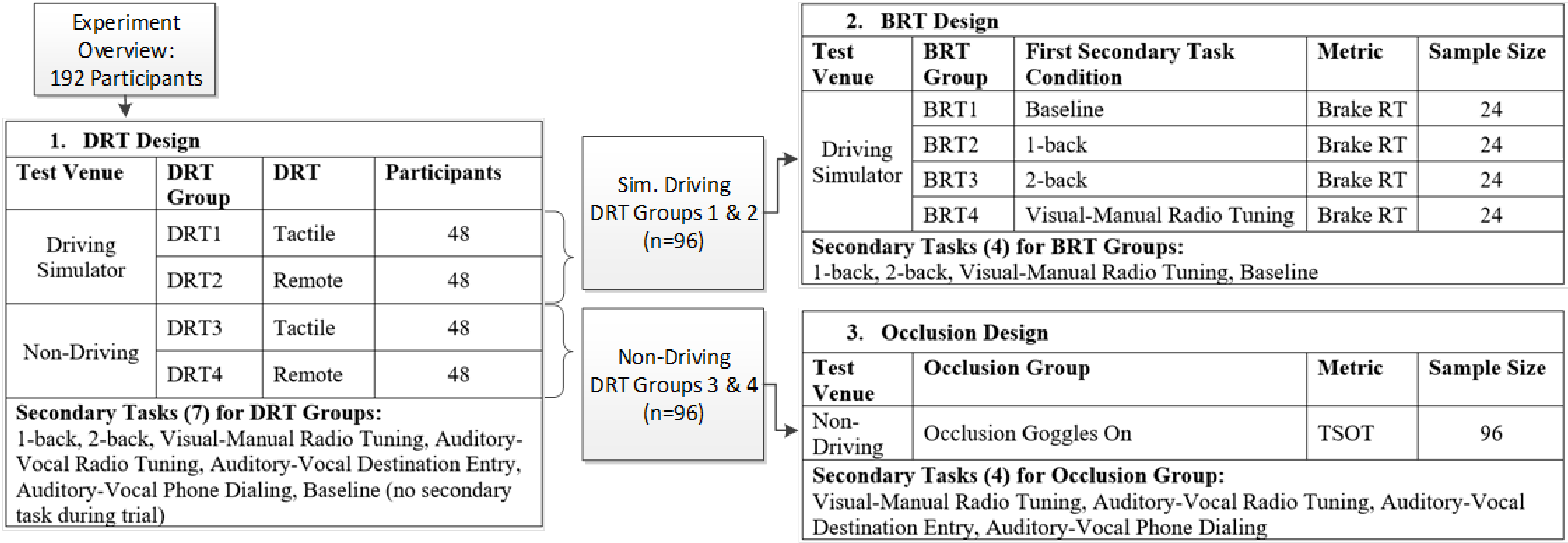
The emergency scenario (defined in Section 1.3) required withholding details of the experimental protocol initially. Specifically, the purpose of this component of the experimental protocol was to obtain data on participants' response time to an unexpected LV braking event. To ensure participants were surprised when the event was presented, there was no information about the event in the informed consent or instructional materials. After the event occurred, the purpose of this component was explained in detail including the need for surprise to elicit an immediate response.

2.1.3 Compensation

Participants were compensated for their participation per guidelines developed by NHTSA (see Appendix F). Compensation consisted of an hourly base pay rate and mileage reimbursement for travel to and from the test facility. The hourly pay rate was $40 per hour.

2.2 Approach

The study objectives were addressed in a single experiment in which test venue (i.e., simulator or non-driving venue) was a between-group factor. One hundred ninety-two drivers participated in the experiment; 96 participants were assigned to each venue. The simulator participants completed two blocks of testing, one related to the DRT and one related to the BRT. The non-driving participants completed two blocks of testing, one related to DRT and one related to Occlusion testing. For DRT testing in both venues, half of the participants used the TDRT and half used the RDRT. The following graphic (Figure 1) is an overview of the participant assignments, with design and descriptions for each component presented in the sections that follow.



Experiment Overview: 192 Participants

1. **DRT Design**

Test Venue	**DRT Group**	**DRT**	**Participants**
Driving Simulator	DRT1	Tactile	48
	DRT2	Remote	48
Non-Driving	DRT3	Tactile	48
	DRT4	Remote	48

Secondary Tasks (7) for DRT Groups:
1-back, 2-back, Visual-Manual Radio Tuning, Auditory-Vocal Radio Tuning, Auditory-Vocal Destination Entry, Auditory-Vocal Phone Dialing, Baseline (no secondary task during trial)

Sim. Driving DRT Groups 1 & 2 (n=96)

Non-Driving DRT Groups 3 & 4 (n=96)

2. **BRT Design**

Test Venue	**BRT Group**	**First Secondary Task Condition**	**Metric**	**Sample Size**
Driving Simulator	BRT1	Baseline	Brake RT	24
	BRT2	1-back	Brake RT	24
	BRT3	2-back	Brake RT	24
	BRT4	Visual-Manual Radio Tuning	Brake RT	24

Secondary Tasks (4) for BRT Groups:
1-back, 2-back, Visual-Manual Radio Tuning, Baseline

3. **Occlusion Design**

Test Venue	**Occlusion Group**	**Metric**	**Sample Size**
Non-Driving	Occlusion Goggles On	TSOT	96

Secondary Tasks (4) for Occlusion Group:
Visual-Manual Radio Tuning, Auditory-Vocal Radio Tuning, Auditory-Vocal Destination Entry, Auditory-Vocal Phone Dialing

Figure 1. Overview of Participant Assignments and Test Conditions

2.2.1 DRT Experimental Design

The DRT experiment was conducted in two test venues, as shown in Figure 1 (1. DRT Design). In the driving-simulator venue, participants performed secondary tasks while performing a car-following task in a fixed-base simulator. In the non-driving venue, participants performed the secondary tasks but had no concurrent driving task. In both venues, the participants also responded to DRT signals, which were presented continuously at 3-5 second intervals to assess the attentional demand associated with the secondary tasks. Participants were assigned to one of four DRT groups shown in Figure 1 (1. DRT Design). In each venue, one group used the TDRT, which used a vibrating stimulus, and one group used the RDRT, which used a visual stimulus. Section 2.3.3 describes the DRT apparatus and Section 2.4 describes the DRT in greater detail.

The design for the DRT experiment was a two-factor mixed design, in which the DRT condition was the between-subjects' variable and the secondary task condition was the within-subjects' variable. Each group of 48 participants was composed of two 24-person samples defined by the NHTSA Driver Distraction Guidelines sample criteria, described previously. All participants performed each of the following secondary tasks in their assigned DRT test condition.

1. N-back easy (1-back)
2. N-back difficult (2-back)
3. Visual-manual radio tuning (V-M radio tuning)
4. Auditory-vocal radio tuning (A-V radio tuning)
5. Auditory-vocal destination entry (destination entry)
6. Auditory-vocal phone dialing (phone dialing)
7. Baseline (no secondary task)

Section 2.6 presents the details of each secondary task.

2.2.2 BRT Experimental Design

In the BRT component, the 96 driving-simulator participants completed a series of four main data collection trials in the simulator, one in each of the following in-vehicle task conditions: 1-back, 2-back, V-M radio tuning, and baseline. Each trial combined car following with seven LV braking events. Among the LV braking events, the first six required drivers to brake in response to LV brake light activation while the seventh was an unexpected crash-imminent event requiring an immediate braking response.

The BRT conditions differed only in terms of the secondary task trial order and the secondary task condition performed first by the driver (i.e., which secondary task was being tested when the first emergency/surprise event occurred), as shown in Figure 1 (2. BRT Design).

Each BRT experimental group was composed of an equal number of participants from each of DRT groups 1 and 2; however, there was no DRT performance in this component. The assignment was intended to ensure there was no bias among BRT groups due to the type of DRT testing preceding the BRT testing. The DRT group assignment resulted in each BRT group having one 24-person NHTSA Driver Distraction Guidelines sample.

2.2.3 Occlusion Experimental Design

In the occlusion experimental component, 96 participants, comprising four 24-person NHTSA Driver Distraction Guidelines sample groups based on age and gender combinations, completed a series of practice and five test trials for each of the four secondary task conditions shown in Figure 1 (3. Occlusion Design).

For half the participants, the occlusion experiment preceded the DRT experiment, while the other half of the participants experienced the DRT experiment first. Each of the four Guidelines sample groups had one person assigned to each of the 24 unique orders of the four in-vehicle tasks.

2.3 Apparatus

All testing was conducted in a laboratory environment at NHTSA's Vehicle Research and Test Center (VRTC) in East Liberty, Ohio.

2.3.1 Test Vehicle

Testing in both venues was conducted with two identical 2011 Ford Explorer SUVs. One SUV was used to control the fixed-base simulator; the second SUV was used for the non-driving test venue. The Ford Explorers' OE SYNC in-vehicle information system was used to present the secondary tasks, except for N-back. Figure 2 presents a close-up view of the steering wheel and console, which consisted of a touch screen above a radio tuning knob.

Figure 2. Steering Wheel and Center Console in 2011 Ford Explorer

2.3.2 Heart Rate Monitor

Heart-rate data was obtained in both test venues. The heart rate monitoring equipment consisted of a Zephyr Bioharness 3 Bluetooth-enabled physiological monitor connected by small wire leads to two 3M Red Dot sensor pads (see Figure 3). The monitor transmitted heart rate data wirelessly to the data acquisition system (DAS) using a Bluetooth connection.

Figure 3. Heart Rate Monitoring Sensors and Wireless Monitor

2.3.3 Detection Response Task Apparatus

2.3.3.1 Tactile Detection Response Task

The TDRT apparatus was created by Red Scientific to meet the requirements of ISO 17488 (ISO, 2014). It consisted of a small electrical vibrator (i.e., tactor) taped on the left shoulder near the clavicle of the participant (Engström, 2010), as shown in Figure 4. The tactor had these specifications: diameter of 10 mm, weight of 1.2 grams, speed of 12,000 rpm, and vibration amplitude of 0.8 G.

Figure 4. Tactile DRT

2.3.3.2 Remote Detection Response Task

The RDRT consisted of a single LED placed remotely in a fixed location near the driver's central field of view (see Figure 5). The RDRT met the requirements of the ISO 17488 (ISO, 2014), which include being red in color with a dominant wavelength of 626 nm, and a 5-mm diameter, however, the ISO recommendation for luminous intensity (i.e., 2 cd) was too bright for use in the laboratory and required adjustment downward.

Figure 5. Remote DRT

2.3.3.3 DRT Response Button

Both DRT variants used a micro-switch (Red Scientific) attached to the participant's left index finger (see Figure 6) for responding to DRT stimuli. The participant tapped his or her left index finger on the steering wheel immediately after detecting the DRT stimulus.

Figure 6. DRT Response Button

2.3.4 Non-Driving Test Venue

For the non-driving test venue, the setup included the following additional components: Occlusion glasses and control hardware, a laptop computer to control the DRTs and secondary task stimulus information, and a DAS to record DRT data, occlusion data, heart rate data, and video from multiple camera locations.

2.3.4.1 Occlusion Goggles

Occlusion goggles were used in the non-driving venue. They are worn like regular glasses but have lenses that can be made to be either transparent or opaque (see Figure 7). The glasses were connected to a computer so that the lenses could be made to alternate between the two conditions.

Figure 7. Occlusion Goggles in Transparent and Opaque Condition

2.3.5 Driving Simulator Test Venue

The fixed-base driving simulator used the National Advanced Driving Simulator miniSim PC-based driving simulation software developed at the University of Iowa. The driving simulator was controlled using the OE steering wheel and pedals of the stationary test vehicle. Figure 8 presents a drawing of the simulator enclosure with the dimensions and layout of the vehicle and equipment inside.

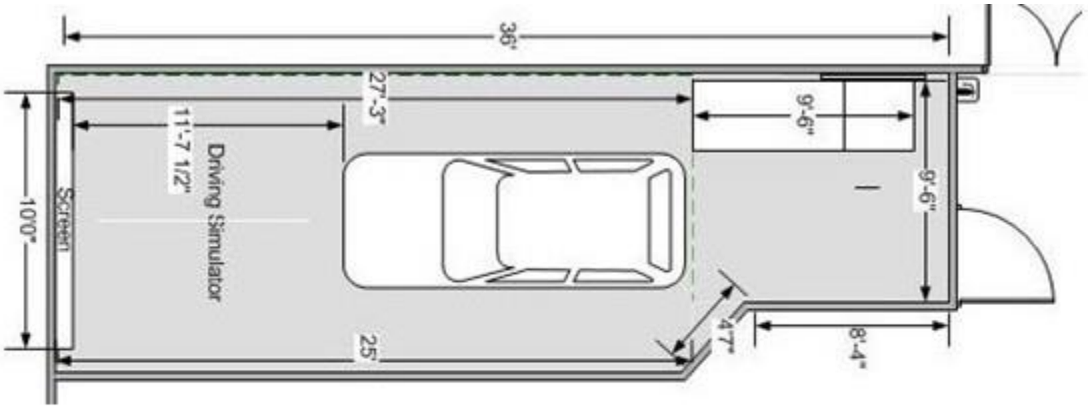


Figure 8. Dimensions and Basic Layout of Simulator Environment

The enclosure was constructed using materials from two portable canopies made of interlocking aluminum poles with tarps covering the roof and sides. Insulated partitions used for walls and additional insulation added to the roof reduce ambient light and provide effective sound proofing for the enclosure. Experimenters and simulator control equipment were located on a bench positioned behind the vehicle on the passenger side (see Figure 8). Experimenters and participants communicated through the vehicle’s open windows and a speaker and microphone system.

In addition to the production vehicle, components of the fixed-base simulator included a (NADS) miniSim computer, a ceiling-mounted digital projector (1024 x 768) positioned above the vehicle, and a forward projection screen (10 x 10 ft), which displayed the roadway image. The screen was located approximately 176 inches in front of an average driver’s eye point. The simulator scenario was programmed using the NADS Interactive Scenario Authoring Tool software. The driving scenario is shown in Figure 9. Simulated road and engine noise, generated by the simulator control software, were presented in speakers located outside the vehicle.

Figure 9. NHTSA Distraction Driver Guidelines Driving Scenario (NADS ISAT Version)

Sensors that recorded steering, accelerator and brake inputs were attached temporarily to the test vehicle. Specifically, a bracket was developed to couple either front wheel of the test vehicle to a turn plate on the ground while the vehicle wheels are off the ground. The test vehicle was supported by jack stands. The bracket and turn plate assembly mounted to the front wheel provided steering inputs to the driving simulator when the participant moved the steering wheel, allowing the simulator to run without the vehicle being turned on to activate the power steering. Having the wheels off the ground allowed the steering wheel to move relatively freely, providing a feel similar to on-road driving. The simulator computer recorded steering wheel, brake, and throttle position inputs. Figure 10 shows the bracket assembly connected to the right front wheel of the study vehicle with the tire removed.

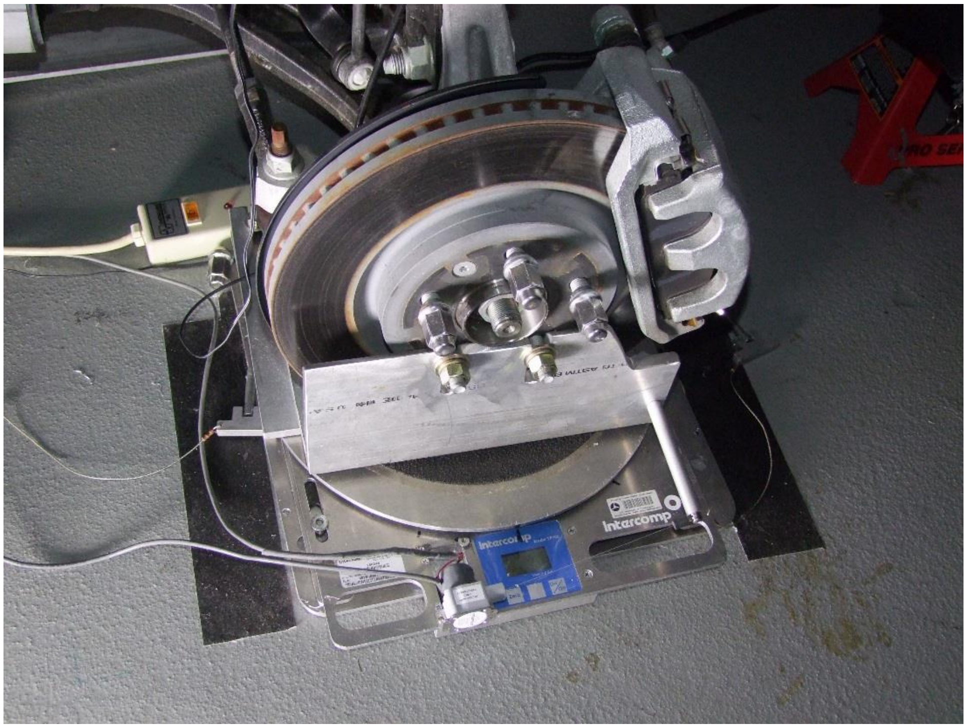


Figure 10. Connection to Test Vehicle for Recording Steering Wheel Movement

A separate computer-controlled data acquisition; it provided time synchronization for all data, which included DRT and BRT data, video from multiple camera locations inside the vehicle, miniSim driving data, and heart rate data. Control software, running in the same computer, controlled stimulus presentation, including DRT stimuli and secondary task instructions and stimuli. It also computed performance summary measures, which allowed the experimenter to monitor the experimental trials and provide feedback to participants following each trial.

Additional details on the laboratory setup are presented in Ranney, Baldwin, Parmer, Domeyer, Martin and Mazzae (2011). Descriptions of the apparatus used to present the DRT tasks are presented in the following sections. BRT tasks were presented via the driving simulator.

2.3.5.1 Head-Mounted Eye Tracker

Dikablis binocular professional head-mounted eye-tracking glasses were used to measure and record eye-glance behavior in the driving simulator (see Figure 11). The eye-tracking glasses were manufactured by Ergoneers of North America, Inc. (2015). Ergoneers develops, manufactures, and distributes measurement and analysis systems for behavioral research, including D-Lab, software for capturing and analyzing human behavior.


Figure 11. Dikablis Head-Mounted Eye Tracker

2.4 Detection Response Task

The DRT was performed continuously while participants completed in-vehicle tasks in both venues. The DRT stimuli, either tactile or visual, were presented at randomly selected intervals between 3 and 5 seconds. Presentation was programmed to ensure that the first stimulus was presented within 3 to 5 s of the onset of each secondary task, such that at least one stimulus occurred during each of the short duration tasks performed. Each stimulus remained activated for 1 s, unless the response button was pressed before then, in which the press of the response button would extinguish the stimulus. A valid response was one that occurred after the stimulus was activated for at least 100 ms, but before 2.5 s of time passed after stimulus onset.

2.5 Driving Tasks

Three driving tasks were used in this study. All 96 simulator participants were assigned to either the constant or complex car-following task for DRT testing. These tasks differed in the difficulty of the car-following task. The third driving task combined a hybrid of the car-following task demands from the first two driving tasks with a series of LV braking events created specifically for BRT testing. All 96 simulator participants performed this driving task during BRT testing. Specific details of each driving task are presented below.

2.5.1 Constant Car-Following Task

This driving task and supporting rationale were specified in the NHTSA Driver Distraction Guidelines. It consisted of a simple car-following task on a straight 4-lane undivided roadway with no intersections. The lead vehicle maintained a constant speed of 50 mph, which was the posted and instructed speed limit. The driving task required participants to maintain a constant following distance of approximately 70 m (220 ft) behind the lead vehicle. Car-following performance feedback was provided to the test participant after each trial. Other than the lead vehicle, there was no traffic present.

2.5.2 Complex Car-Following Task

This driving task consisted of a complex car-following task on a straight 4-lane undivided roadway with no intersections. The lead vehicle varied its speed based on a complex function created to combine several sine waves selected based on levels of acceleration and deceleration typically experienced (see Appendix G). The average speed was approximately 50 mph. This driving task required participants to maintain a constant following distance of approximately 37 m (120 feet) behind the lead vehicle. If the following distance was outside the specified acceptable range (i.e., 60 to 180 feet) for more than 5 s, a warning tone was activated intermittently until the following distance was within the acceptable range. Car-following performance feedback was provided after each trial. Other than the lead vehicle, there was no traffic present.

2.5.3 Car Following with Lead-Vehicle Braking Events

This driving task consisted of a simple car-following task on a straight, four-lane, undivided roadway with no intersections. The LV maintained a constant speed of 50 mph, except during the seven LV braking events that occurred during each trial. If the following distance was beyond the specified acceptable maximum distance (150 ft) for more than 5 s, a warning tone was activated intermittently until the following distance was less than the acceptable maximum distance. Car-following performance feedback was provided after each trial. Other than the lead vehicle, there was no traffic present.

This task required participants to maintain a constant following distance of approximately 37 m (120 ft) and respond to any LV brake light activation by quickly pressing the brake pedal. The seven LV braking events occurred at intervals of approximately 30 s. For the first six LV braking events, the LV brake lights were activated to indicate braking. During these events, the LV slowed to 40 mph at a rate of 4 m/s^2 for a period of 2 s but did not stop. The LV then accelerated to 50 mph and maintained that speed until the next event. The participants were instructed to press the brake pedal quickly when they noticed LV brake light activation. These braking events were referred to as expected LV braking events. Each trial ended with a seventh, emergency LV braking event, which involved the lead vehicle stopping at a rate of 5 m/s^2 with no brake light activation. This event required an immediate and sustained braking response to avoid collision. After 8 s, including stopping and stopped time, the LV accelerated to 50 mph as the trial ended.

The braking events were positioned in the driving scenario to coincide with the performance of V-M radio tuning task instances; odd numbered braking events occurred during radio tuning task instances, while even numbered events occurred between radio tuning task instances. Such event matching was not relevant for the continuous in-vehicle tasks (1-back and 2-back), which required continuous performance throughout the drive. Thus, one scenario worked for all four in-vehicle tasks shown in Figure 1 (2. BRT Design).

2.6 Secondary Tasks

The secondary task conditions used in the DRT testing are summarized in Table 1. The column entries indicate whether the tasks have A-V and/or V-M demands.

Table 1. Secondary Task Conditions

Task	**Auditory-Vocal (A-V)**	**Visual-Manual (V-M)**
N-back (1-back)	Yes	No
N-back (2-back)	Yes	No
V-M Radio Tuning	No	Yes
A-V Radio Tuning	Yes	Yes
Destination Entry	Yes	Yes
Phone Dialing	Yes	Yes
Baseline (no secondary task)	N/A	N/A

The N-back task conditions had no V-M components. The V-M radio tuning task had no A-V demands. The remaining three tasks (i.e., A-V radio tuning, destination entry and phone dialing) were primarily auditory-visual, but had some visual-manual components, such as finding and pressing the push-to-talk button on the steering wheel before using a voice command. Details of each secondary task are presented in the following sections.

2.6.1 N-back Task

N-back is a verbal response delayed-digit-recall task in which a participant listens to and repeats a sequence of recorded single digits per one of several specific rules (e.g., 0-back, 1-back, or 2-back) (Mehler, Reimer, & Dusek, 2013). Administration involves presentation of a string of recorded single digits at a predetermined rate (approximately every 2.5 s) over a data collection interval of several minutes. The participant is asked to respond after each presentation with either the same digit that was just presented (0-back, easy condition) or with the digit that was previously presented (1-back, low difficulty condition) or with the digit that was presented before the previously presented digit (2-back, moderate difficulty condition). The sequences for the 0-back, 1-back, and 2-back conditions are shown in Table 2.

Table 2. N-back Stimulus and Response Sequence by Task Difficulty

Task	**Digit presented**	**3**	**2**	**6**	**7**	**1**
0-back	Correct response	3	2	6	7	1
1-back	Correct response	-	3	2	6	7
2-back	Correct response	-	-	3	2	6

In this table, the task sequence over time moves from left to right. The experimenter’s presentation, which is the same for all conditions in this example, is shown in the top row. The correct response to be said aloud by the participant is presented in the rows beneath the digit presented. The 1-back and 2-back conditions differ from the 0-back condition in that they place a greater burden on working memory, which is a main component of cognitive demand. The 0-back condition is shown here to create a better understanding of the N-back task; however, the 0-back condition was not used in this study.

2.6.2 Radio Tuning

Two versions of radio tuning were used, including a V-M version and an A-V version. Both versions were performed using the Ford SYNC in-vehicle information system in the test vehicle. The following sections present the sequences of actions required for the respective tasks types.

2.6.2.1 Visual-Manual Radio Tuning

Each task instance began with the playing of a pre-recorded instruction that consisted of the frequency band (AM, FM), the frequency, and the word “Begin.” The V-M radio tuning task sequence of actions were as follows:

- Audio file plays: “AM 530. Begin.”
- Participant selects frequency band by pressing AM or FM button on the screen [1]. See Figure 12 for an example of the interface.
- Participant uses “up” or “down” tuning button [2] below the display to scroll to the correct frequency.
- Participant visually confirms the band and frequency [3] are correct on the radio screen.
- Participant says: “Done.”

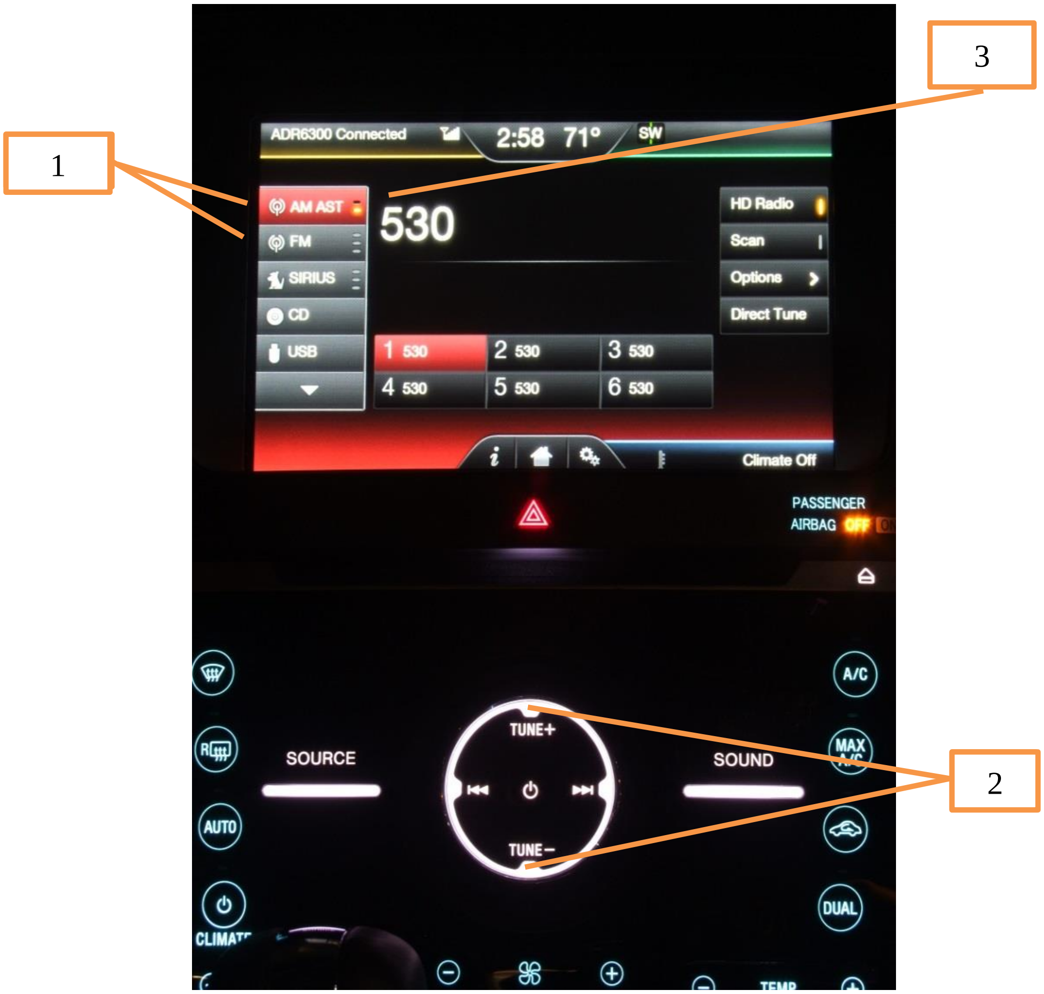


Figure 12. Visual-Manual Radio Tuning Interface

2.6.2.2 Auditory-Vocal Radio Tuning

Each task instance began with the playing of a pre-recorded instruction that consisted of the frequency band (AM, FM), the frequency, and the word “Begin.” The A-V radio tuning task sequence of actions were as follows.

- Audio file plays: “FM 92.3. Begin.”
- Participant presses the “Push to Talk” button on the steering wheel [1]. See Figure 13.
- Vehicle system audio says: “Please say a command.”
- Participant says: “FM 92.3,” when device is in “listening” mode [2]. Figure 14.
- Vehicle system audio says: “Tuning to FM 92.3.”
- Radio displays FM 92.3 [3]. See Figure 15.Participant says: “Done.”

Figure 13. Push-to-Talk Button on Steering Wheel

Say a command.

You can also say:

Figure 14. Interface With Command Options: System Listening for Voice Command

Figure 15. Interface After System Receives Radio Tuning Command

2.6.3 Phone Dialing

Phone dialing was performed using voice commands with the SYNC in-vehicle information system in the test vehicle, which was connected to a wireless phone via Bluetooth. Each task instance began with the playing of a pre-recorded instruction that consisted of a 10-digit phone number and the word "Begin." The phone dialing task sequence of actions were as follows.

- Audio file plays: "9-3-7-5-5-5-1-2-1-2. Begin."
- Participant presses the "Push to Talk" button on the steering wheel [1]. See Figure 13.
- Vehicle system audio says: "Please say a command."
- Participant says: "Dial." when device is in "listening" mode [2]. See Figure 14.
- Vehicle system audio says: "Start saying a phone number" [3]. See Figure 16.
- Participant says: "9-3-7-5-5-5-1-2-1-2."
- Vehicle system audio says: "9-3-7-5-5-5-1-2-1-2. Say Dial, Delete, or Continue speaking the digits" [4]. See Figure 17.
- Participant says: "Dial" [5]. See Figure 18.
- Vehicle system audio says: "Dialing."
- Audible ringing precedes appearance of red receiver (to end call) on the display [6], upon which participant presses the "End Call" icon on the screen. See Figure 19.

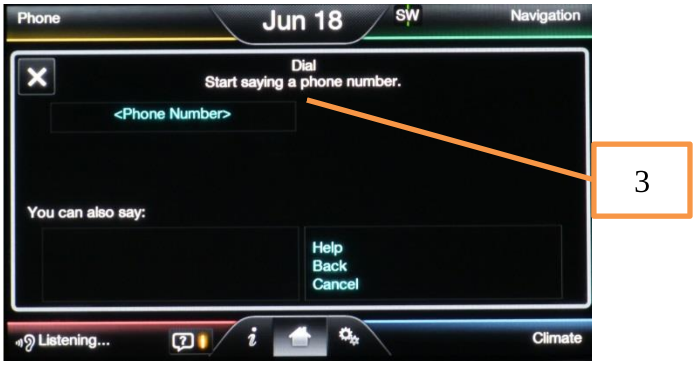

Dial
Start saying a phone number.

Figure 16. Interface After System Receives Dial Command

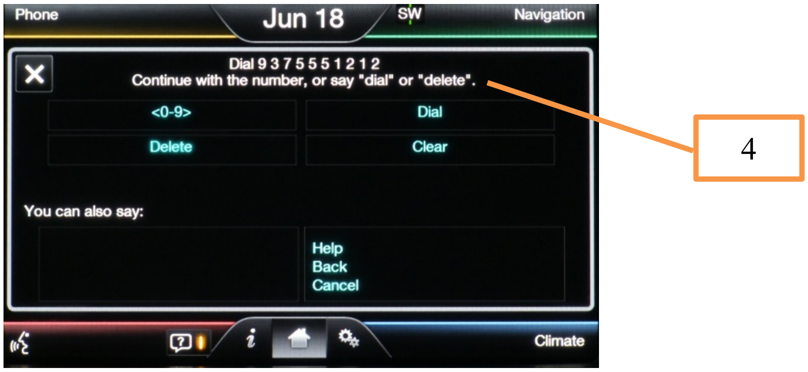

Dial 9 3 7 5 5 5 1 2 1 2
Continue with the number, or say "dial" or "delete".

Figure 17. Interface After System Receives Phone Number

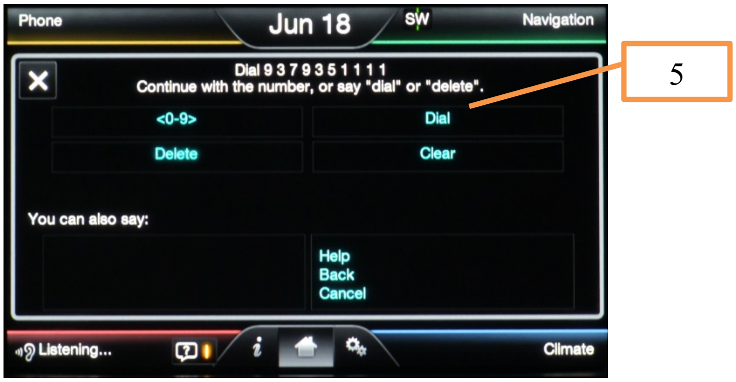

Dial 9 3 7 9 3 5 1 1 1 1
Continue with the number, or say "dial" or "delete".

Figure 18. Interface With Command Options: System Listening for “Dial” Command

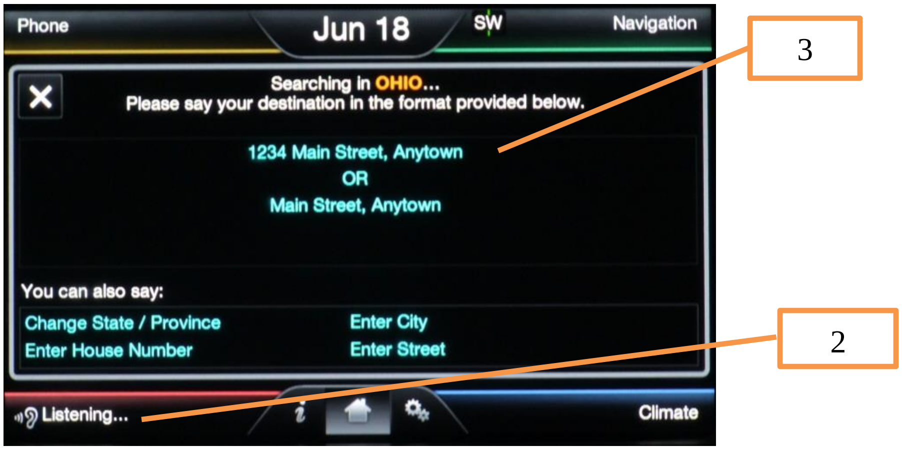
Figure 19. Interface With “End Call” Option

2.6.4 Destination Entry

Destination entry was performed using voice commands with the SYNC in-vehicle information system in the test vehicle. Each task instance began with the playing of a pre-recorded instruction that consisted of an address (house number, street name and city name) and the word “Begin.” The destination entry task sequence of actions was as follows:

- Audio file plays: “1-7-6 Walker Street, Tiffin. Begin.”
- Participant presses the “Push to Talk” button on the steering wheel [1]. See Figure 13.

- Vehicle system audio says: “Please say a command.”
- Participant says: “Find an Address” when device is in listening mode [2]. See Figure 20.
- Vehicle system audio says: “In Ohio. Say the street address including the city” [3]. See Figure 20.
- Participant says: “1-7-6 Walker Street, Tiffin.”
- Vehicle system displays map with address entered [4] and says: “When ready, press the Voice button and then say Set as destination, Set as waypoint, or Change something.” See Figure 21.
- Participant verifies the address is correct.
- Participant says: “Done.”
- Participant presses the “Home” icon on the screen.


Searching in OHIO...
Please say your destination in the format provided below.

1234 Main Street, Anytown
OR
Main Street, Anytown

You can also say:

Figure 20. Interface Listening for an Address in Ohio


176 WALKER ST, TIFFIN, OH

Figure 21. Interface Showing Map Screen With Destination Entered

2.6.5 Secondary Task Timing

Stimuli for secondary tasks were presented at fixed intervals to control the level of task demand. N-back stimuli were presented every 2.5 s to achieve six groups of 10 numbers during each main trial. Task intervals for the tasks performed using the in-vehicle information system are presented in Table 3. Using a V-M radio tuning trial as an example, a participant would receive a new radio tuning stimulus every 36 s until the trial was over. With these timing intervals, participants could complete the clear majority of the tasks in the time frame allotted and have a bit of down time before the next stimuli was presented in a trial.

Table 3. Secondary Task Presentation Intervals by Test Venue (seconds)

Secondary Task	**Number of Secondary Tasks per DRT Trial**	**Test Venue**	
		Non-Driving	**Simulator**
Phone Dialing	4	45	45
Destination Entry	3	60	60
A-V Radio Tuning	6	30	30
V-M Radio Tuning	5	36	36

2.7 Procedures

The procedure consisted of five parts, including: (1) introduction, general instructions and informed consent; (2) training and practice; (3) DRT testing; (4) BRT testing or occlusion testing; and (5) participant debriefing. Each part is discussed in detail below.

2.7.1 Introduction, General Instructions and Informed Consent

Selected participants were assigned to one of two test venues and scheduled individually for a single session of approximately 4-5 hr. The Participant Information Summary and Confidential Information Form (Appendix C), referred to as the informed-consent form was e-mailed to the participants after their session was scheduled. Upon arrival at the test facility, participants were greeted and escorted to a private office or conference room. They were asked to confirm they had received the ICF via e-mail and encouraged to ask questions. For those who had not received the ICF (e.g., last-minute appointments), a hard copy ICF was provided and the participant was given time to review the form. When the participants indicated that they had read the form, one of the co-investigators presented a brief oral summary of the protocol and participation requirements after which the participants were given another opportunity to ask questions. Once all questions had been answered, they were asked to sign an electronic version of the ICF on a tablet computer. Using the same method, they also signed the Information Disclosure form that accompanied the ICF. If a participant refused to sign the ICF, he or she would have been paid for mileage to and from the data collection site and the hourly rate for the time spent at the site. None of the participants refused to sign the ICF. None of the participants declined to participate after being fully informed.

2.7.2 Training and Practice

After consenting to the terms of participation, the participant was given an overview of the experimental protocol and was then escorted to one of two identical experimental vehicles depending on the assigned DRT test protocol (see Figure 1, 1. DRT Design) to begin receiving the appropriate instructions. When seated in the vehicle, the participant was given an overview of the controls and displays and shown how to adjust the seat. Next, the participant was given a description of the task scoring criteria used for performance feedback.

At this point, participants were given instructions based upon which test venue they were assigned to, the driving simulator or the non-driving venue. Participant instructions are presented in Appendix D for both test venues. In the simulator test venue, participants received instructions, training, and practice on the driving task, then the car-following task, then the designated DRT, and then the secondary task. In the non-driving test venue, half of the participants received instructions, training, and practice on the occlusion goggles first, then the designated DRT, and then the secondary task. The other half of the participants received instructions, training, and practice on the designated DRT first, then the occlusion goggles, and then the secondary task. Then, combinations of tasks were practiced together in preparation for a data collection trial on that combination (such as DRT and radio tuning together). After each instruction and training component, the participant was given an opportunity to ask questions about any aspect of the protocol. Participants were regularly offered breaks, which allowed a participant to step out of the vehicle if needed.

2.7.3 DRT Testing

Participants completed all DRT testing in their assigned venue (see Figure 1, 1. DRT Design). DRT testing involved approximately 20 trials. Each trial lasted approximately 3.5 min, with a 3-min continuous data collection interval. Pre-recorded auditory secondary-task stimuli (e.g., instructions for radio tuning or phone-dialing task instances) were presented at predetermined intervals to ensure a consistent level of task demand over the data-collection interval. Additional details for each venue are presented in the following sections.

2.7.3.1 Driving Simulator Test Venue

The participant was instructed to perform the driving tasks (i.e., car following and lane maintenance) plus the specified DRT on each trial. The participant was also given instructions for the designated secondary task. The participant then completed a practice drive with the specified combination of DRT and secondary task. If the participant was comfortable, the data-collection trial was initiated. After each trial, the participant was provided performance feedback. The experimenter then described the next trial and secondary task.

The experimenters were positioned at a control station behind the vehicle during data collection. Communication with the participant was accomplished via direct interaction through the vehicle’s open windows and using a speaker and microphone system. Training, practice, and testing was completed in the order assigned to each participant.

2.7.3.2 Non-Driving Test Venue

Half of the participants in the non-driving venue received the occlusion protocol first, while the other half received the DRT protocol first. In both protocols, the pattern of instruction, training and practice were the same. In one protocol, the participant instructions were based on learning how to respond to the designated DRT, while in the other protocol, the participant instructions were based on how the occlusion goggles worked.

The participant was instructed to perform the specified DRT or wear the occlusion goggles on each trial. The participant was also given instructions for the designated secondary task. The participant then completed a practice trial with the specified combination of DRT (or occlusion) and secondary task. If the participant was comfortable, the data-collection trial was initiated. After each trial, the participant was provided performance feedback. The experimenter then described the next trial and secondary task.

The experimenters were positioned at a control station behind the vehicle during data collection. Communication with the participant was accomplished via direct interaction through the vehicle's open windows and by using a speaker and microphone system. Training, practice, and testing were completed in the order assigned to each participant.

2.7.4 BRT Testing or Occlusion Testing

All participants completed DRT testing in their assigned test venue. Participants assigned to the non-driving test venue also completed occlusion testing, as described above.

Participants assigned to the simulator venue completed DRT testing before starting BRT testing. BRT testing always occurred after DRT testing was complete such that participants were already very familiar with driving the simulator. The pattern for instruction, training and practice was the same as that used for the DRT testing.

Following a short break from DRT testing, the participant was escorted back into the driving simulator for the brake-response test component training and practice. No DRT was used during this component of the testing. Participants got to practice the combination of braking events and secondary tasks before each data collection trial. Each data-collection trial involved the surprise event occurring after the six braking events, approximately 3 min into the data-collection drive.

The trial terminated shortly after the surprise event. After the first surprise event trial, the experimenter explained the need to withhold information about the event to ensure the surprise response. Once all four trials were completed, the participant was asked to complete the simulator sickness questionnaire (Appendix E) to determine if additional rest was required before being allowed to drive home.

2.7.5 Participant Debriefing

At the completion of data collection, the participant exited the vehicle and proceeded to the designated office or conference room. The participant's pay was calculated based on two amounts, hourly base pay for participation and mileage for travel to and from the test site. Payment was made to the participant in person at the completion of the session. The experimenter answered questions posed by the participant and escorted the participant to his or her personal vehicle.

3.0 RESULTS

Analyses were conducted to address the eight study objectives. The results are presented in the following five sections, addressing DRT, BRT, occlusion, eye-glance and heart-rate analyses, respectively.

3.1 DRT Analyses

This section presents results pertaining to study objectives one, two, four and five.

3.1.1 DRT Response Performance

DRT performance was assessed with two metrics, response time and proportion of correct responses (accuracy). Data points used in these analyses represent mean RT or detection accuracy performance for each participant in a given task condition. Performance was summarized over the intervals during which the secondary task was performed. Data was combined to create a composite of DRT data from at least 2 min of time during which the participant was performing the secondary task. DRT events occurred at approximately 4-s intervals, which provides approximately 30 targets in a 2-min summary interval. Thus, each data point in the following analyses represents the mean RT for all targets detected correctly within this interval, while accuracy represents the proportion of these (approximately 30) targets that were correctly detected.

The DRT analysis had two primary objectives; first, to determine whether differences exist between driving and non-driving test venues; and second, to determine whether differences exist between the remote and tactile versions of DRT. Box plots of distributions are presented for each metric in each test venue. Figure 22 presents distributions of simulator DRT RT by task and DRT condition. Boxes enclose the middle 50 percent of each distribution. Horizontal lines inside each box represent median values; symbols inside each box are positioned at the distribution means.

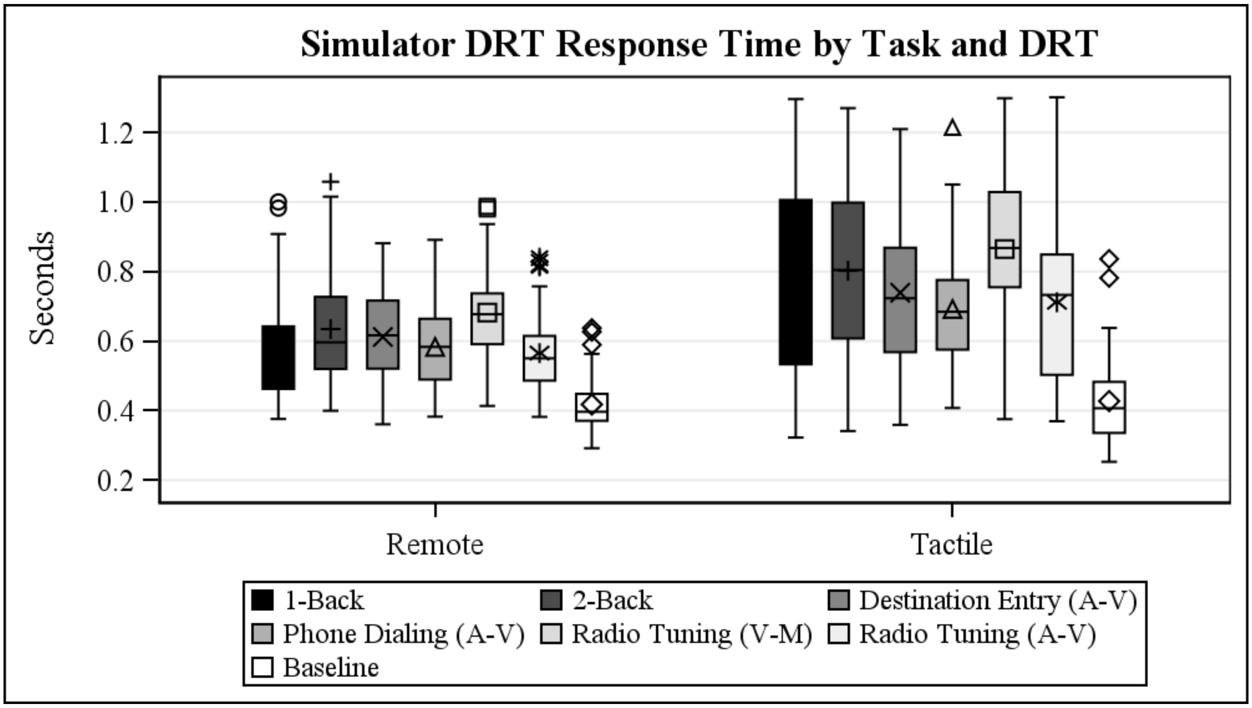


Figure 22. Simulator DRT Response Time by Task and DRT Conditions

Except for the baseline condition, responses to tactile targets were consistently slower than responses to visual targets in the simulator. Tactile mean RTs were also consistently more variable than remote mean RTs as indicated by the longer boxes and ranges associated with the distributions of RT mean values in that condition. Distributions of simulator DRT accuracy by task and DRT condition are presented in Figure 23.

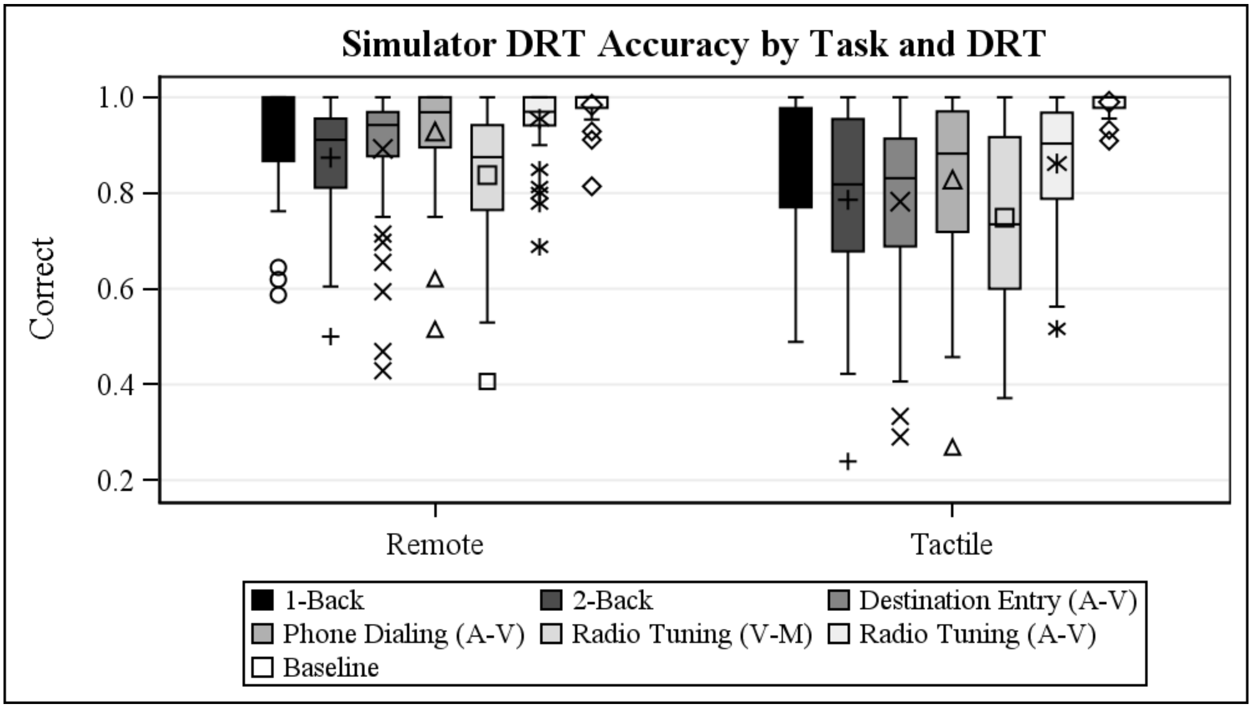


Figure 23. Simulator DRT Accuracy by Task and DRT Conditions

Response acccuracy was generally higher for remote (i.e., visual) targets than for tactile targets in the simulator test venue. However, the RDRT accuracy metric appears constrained by the ceiling of perfect performance for several task conditions. The increased variability apparent in the TDRT RT distributions is also apparent in the TDRT accuracy distribution. Thus, response accuracy was more variable among participants in the Tactile condition than among those in the RDRT condition.

We explored the relation between RT and accuracy to determine whether a speed-acuracy tradeoff may have been occurrring. The absence of a speed-accuracy tradeoff is an underlying assumption of DRT testing as set forth by the ISO (2014). The scatter plot for the simulator data by DRT type is presented in Figure 24.

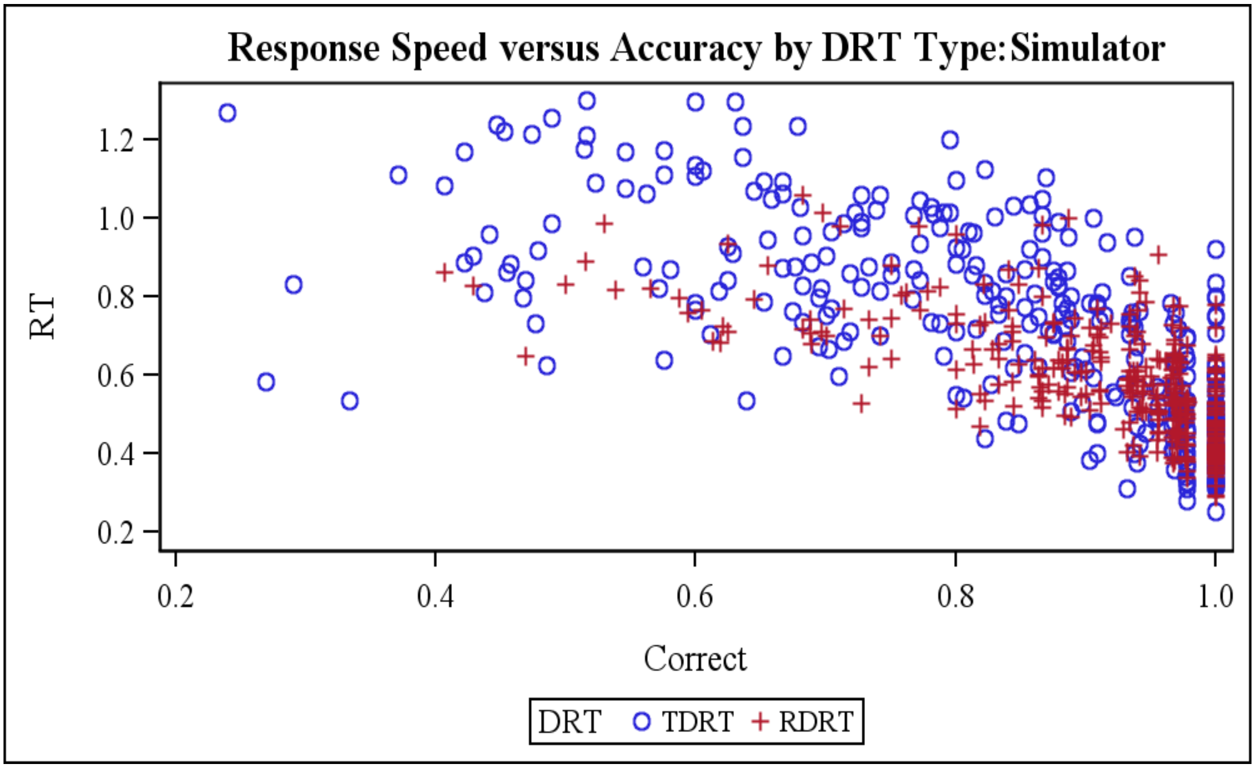


Figure 24. Response Speed Versus Accuracy by DRT Type: Driving Simulator

Spearman rank correlation values between RT and accuracy were *R* = -0.71 for RDRT and *R* = -0.76 for TDRT. These values indicate that participants were not trading speed for accuracy; rather, high accuracy was more likely associated with faster response speeds. The scatter plot does reveal the ceiling effect associated with the limit of perfect performance.

Figure 25 presents the distributions of RT means in the non-driving venue by DRT type and task condition.

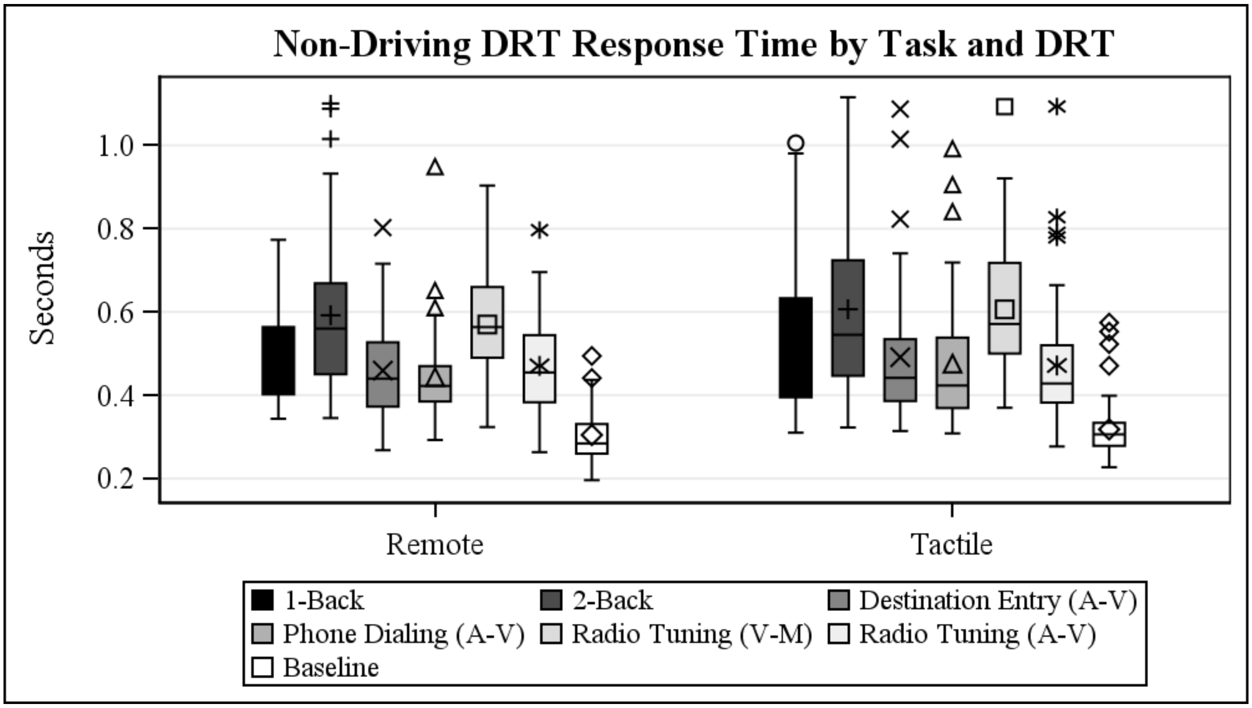


Figure 25. Non-Driving DRT Response Time by Task and DRT Conditions

The differences between DRT conditions among mean RT values in the non-driving venue appear considerably smaller than the corresponding differences in the driving simulator. Specifically, mean RT values in the TDRT condition do not appear to be consistently greater than those in the RDRT condition. And, while the TDRT distributions reveal greater variability among RT mean values than the RDRT distribution, the differences are considerably smaller in the non-driving venue than for the simulator data.

Figure 26 presents the distribution of DRT accuracy by DRT and task condition for the non-driving venue.

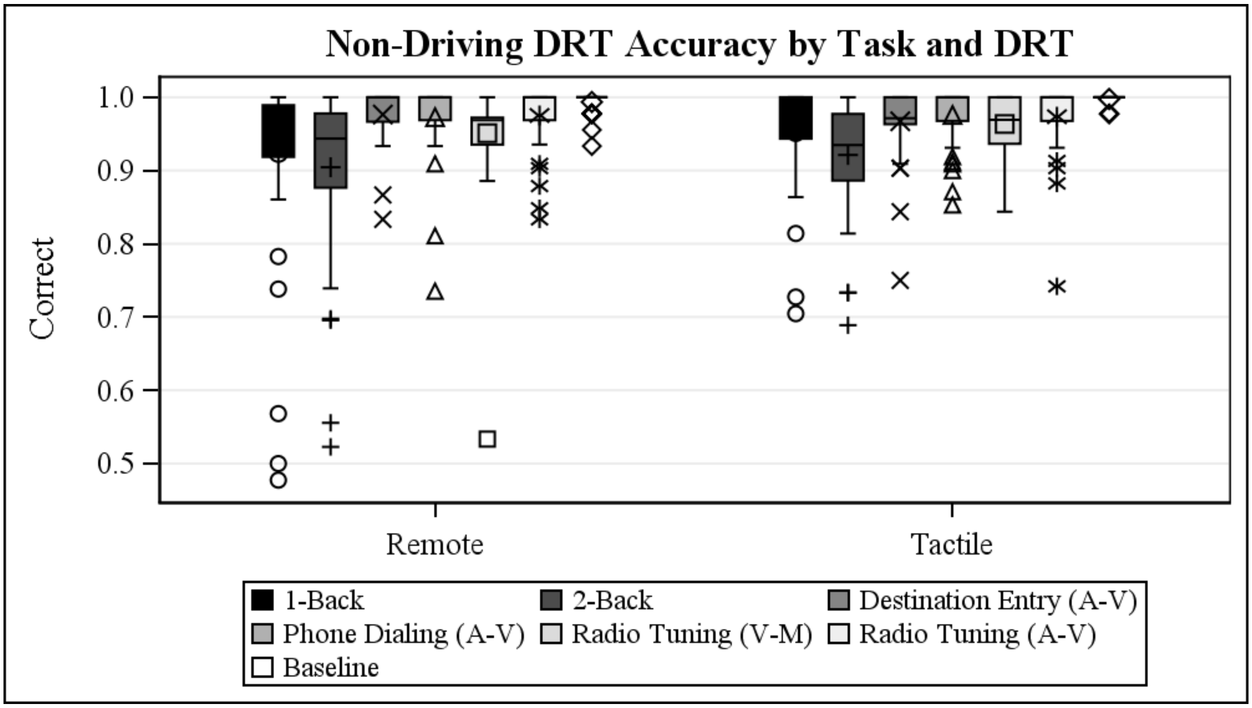


Figure 26. Non-Driving DRT Accuracy by Task and DRT Conditions

With no driving task, DRT accuracy in the non-driving condition is generally higher than in the simulator. In particular, DRT accuracy in the non-driving venue hits the ceiling of perfect performance for most task conditions. There appear to be only minor differences between the two DRT conditions. Unlike the simulator data, which showed reduced accuracy for tactile targets, the detection accuracy appears similar in both DRT conditions across task conditions in the non-driving venue.

Figure 27 presents the scatter plot of the two DRT metrics by DRT type for the non-driving venue. Spearman rank correlation values between RT and accuracy were R = -0.56 for RDRT and R = -0.54 for TDRT. Although the correlations are weaker than those associated with the simulator venue, these values indicate that participants were not trading speed for accuracy; rather, higher accuracy was more likely associated with faster response speeds. As shown in Figure 27, the relation between speed and accuracy appears truncated by the high proportion of trials with perfect performance. This pattern was evident to a lesser extent for the simulator data (Figure 24). The higher proportion of perfect trials in the non-driving venue was likely the main cause of smaller correlation values observed in this venue.

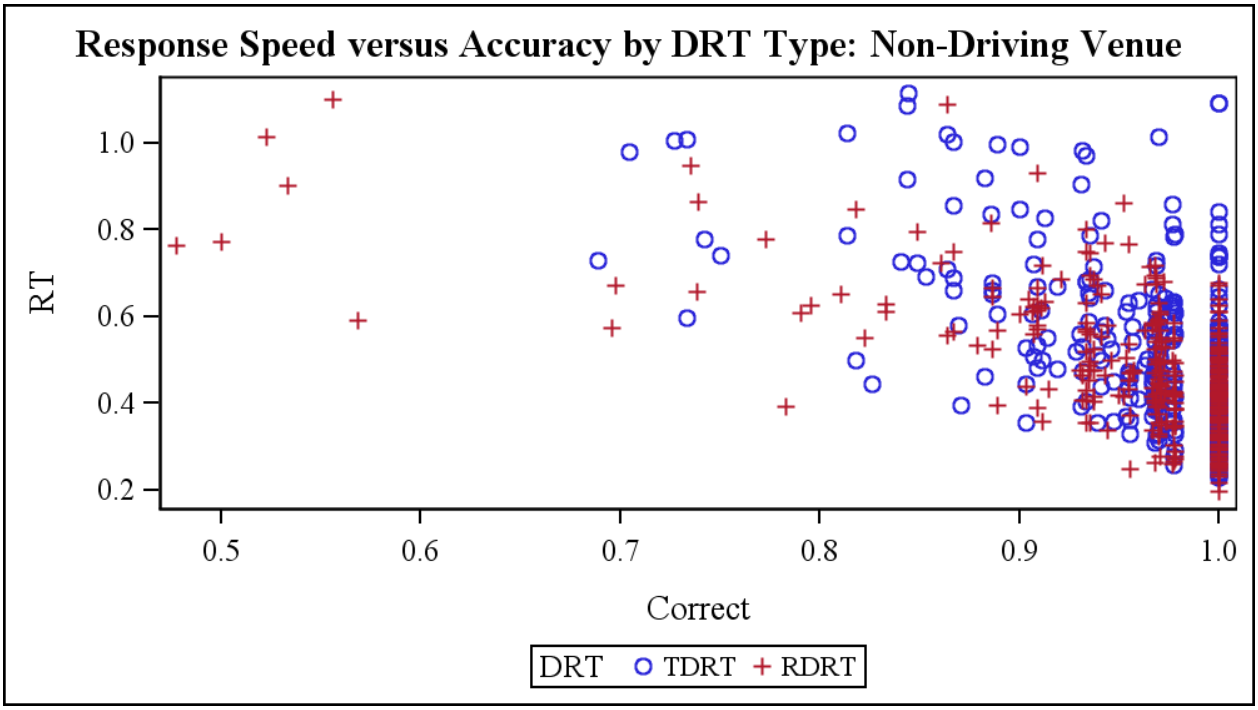


Figure 27. Response Speed Versus Accuracy by DRT Type: Non-Driving Venue

3.1.2 DRT Statistical Testing

Statistical analyses were conducted to assess the effects of independent variables on DRT RT and accuracy. These analyses were conducted first for each venue separately and then together to examine differences between venues.

3.1.2.1 Simulator DRT RT

Because neither response variable was normally distributed, Proc Glimmix (SAS V. 9.4) was used to perform the analyses. RT was modeled using a gamma distribution. Independent variables included DRT type, age group, and task condition. The baseline DRT RT was included as a covariate. Car-following speed was initially included in the model, but was removed when it was revealed that there were no differences associated with this variable. Table 4 presents the ANOVA summary table for fixed effects for this analysis.

Table 4. ANOVA Fixed Effects: DRT Simulator Response Time

Type III Tests of Fixed Effects				
Effect	**Numerator DF**	**Denominator DF**	**F Value**	**P > F**
DRT	1	87.01	28.71	<.0001
Age Group	3	87	6.41	0.0006
Baseline RT	1	87.04	34.98	<.0001
Task	5	453.1	25.58	<.0001
DRT*Task	5	453.1	2.14	0.0593
DRT*Age Group	3	87.01	0.42	0.7384
Age Group*Task	15	453.1	2.27	0.0043

The DRT x Task Condition interaction was marginal but did not attain statistical significance. However, a set of planned comparison tests was conducted to examine differences between the two DRT conditions. The comparisons examined differences between each of the four in-vehicle tasks and the two benchmark conditions (i.e., 1-back and 2-back) in each DRT condition. The results of these comparisons are presented in Table 5.

Table 5. Pairwise Differences Between Task Conditions by DRT Condition (Simulator RT)

	RDRT			**TDRT**			**Consistent**
Comparison	**DF**	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	453	-3.79	0.0025	453	-1.54	0.789	No
1-back vs. Destination Entry	453	-2.96	0.0487	453	0.85	0.9959	No*
1-back vs. Phone Dialing	453	-1.32	0.9009	453	2.81	0.0724	No*
1-back vs. V-M Radio	453	-6.84	<.0001	453.4	-4.36	0.0004	Yes
1-back vs. A-V Radio	453	-0.21	1	453	2.49	0.1656	No*
2-back vs. Destination Entry	453	0.83	0.9964	453	2.38	0.2171	Yes
2-back vs. Phone Dialing	453	2.47	0.1755	453	4.35	0.0004	No
2-back vs. V-M Radio	453	-3.05	0.0382	453.4	-2.85	0.0662	No
2-back vs. A-V Radio	453	3.58	0.0054	453	4.03	0.0012	Yes

*Direction of differences not consistent

Six of the nine planned comparisons had inconsistent outcomes across DRT types, while three had consistent results. Three of the comparisons had differences in different directions. These included the following differences: 1-back versus destination entry; 1-back versus phone Dialing, and 1-back versus A-V radio tuning, although this latter difference was not significant for both DRT conditions. Of interest is the difference between the two benchmark conditions (i.e., 1-back and 2-back), which has been consistently found in previous studies. In our simulator data, the RDRT demonstrated the expected difference between these two conditions, while the TDRT did not. Because the DRT stimuli use different modalities (i.e., visual for RDRT; tactile for TDRT), one might expect differences between DRT conditions to exist for comparisons involving the V-M radio tuning task for which the associated visual demands could affect the DRTs differently.

The pattern of results for the 2-back versus V-M radio tuning comparison across DRT conditions appears consistent with this expectation; however, the difference between outcomes noted in the table for this comparison is marginal. Means for the simulator DRT RTs by Task and DRT condition are presented in Figure 28.

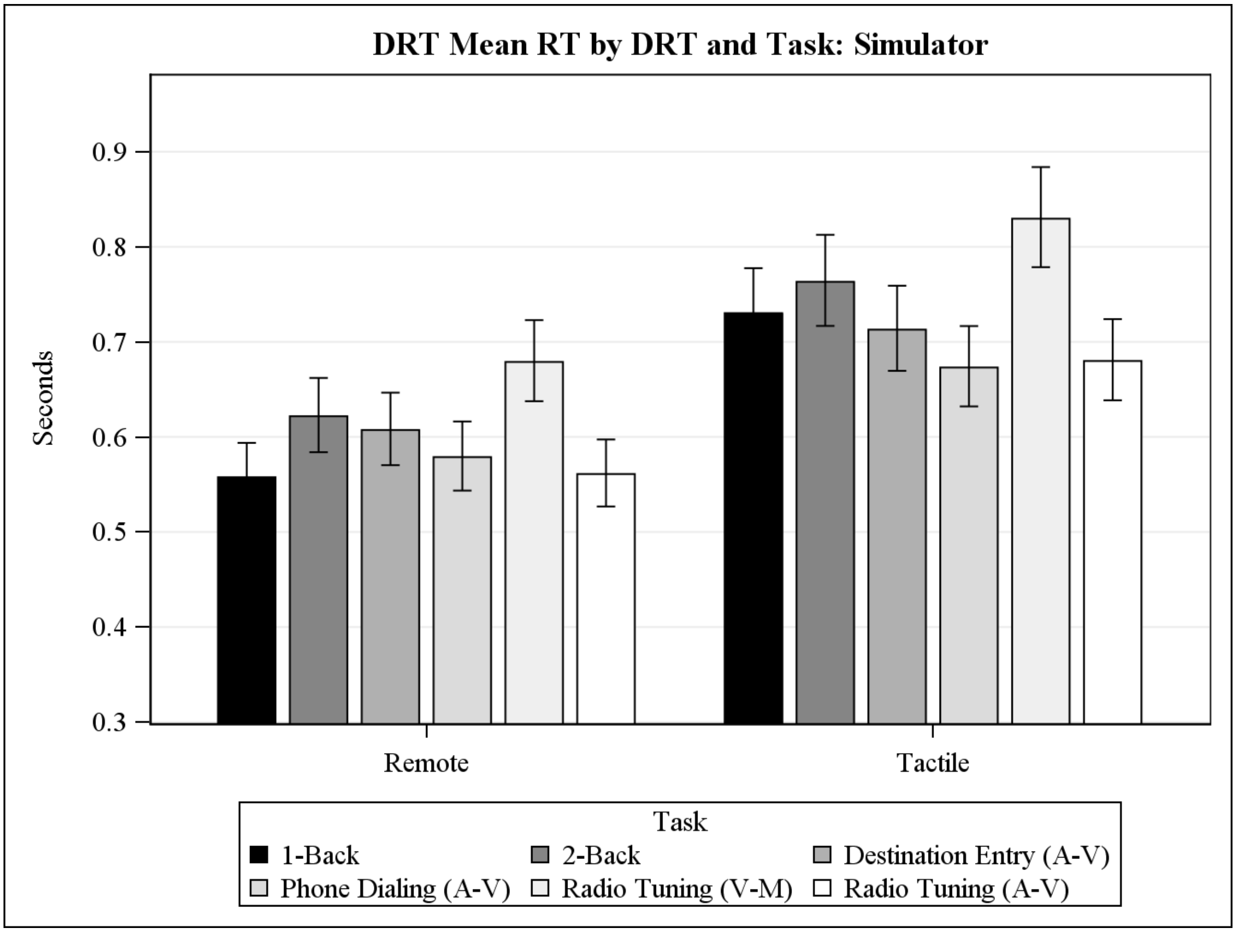


Figure 28. DRT Mean RT by DRT Condition and Task

3.1.2.2 Simulator DRT Accuracy

For this analysis, DRT accuracy, defined as the proportion of correct responses while performing the task, was modeled as a binomial variable since the response variable for each trial had only two possible outcomes, detect or fail-to-detect (i.e., miss). Car-following speed was removed from this model as with RT. The independent variables included in this analysis were the same as for DRT RT.

Table 6 presents the ANOVA summary table for fixed effects for this analysis.

Table 6. ANOVA Fixed Effects: DRT Simulator Response Accuracy

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
DRT	1	83.58	20.40	<.0001
Age Group	3	83.27	11.28	<.0001
Baseline Accuracy	1	78.09	4.54	0.0362
Task	5	540	48.27	<.0001
DRT*Task	5	540	5.15	0.0001
DRT*Age Group	3	81.87	0.22	0.8789
Age Group*Task	15	540	6.44	<.0001

Main effects of DRT, age group, and task were statistically significant as were the DRT x Task and Age Group x Task interaction effects. Baseline detection accuracy was statistically significant, but apparently much weaker for this metric than for RT, most likely reflecting the fact that individual differences in detection accuracy are weaker than individual differences in response speed.

Table 7 presents the results of planned pairwise comparisons between secondary task conditions in each DRT condition to explore differences in sensitivity between the two DRTs.

Table 7. Pairwise Differences Between Task Conditions by DRT Condition (Simulator Accuracy)

	RDRT			**TDRT**			**Consistent**
	DF	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	540	4.22	0.0007	540	5.19	<.0001	Yes
1-back vs. Destination Entry	540	0.71	0.9993	540	3.94	0.0016	No
1-back vs. Phone Dialing	540	-2.8	0.0751	540	0.25	1	No*
1-back vs. V-M Radio	540	5.87	<.0001	540	7.37	<.0001	Yes
1-back vs. A-V Radio	540	-5.9	<.0001	540	-2.24	0.3019	No
2-back vs. Destination Entry	540	-3.19	0.0228	540	-0.89	0.995	No
2-back vs. Phone Dialing	540	-6.42	<.0001	540	-4.43	0.0002	Yes
2-back vs. V-M Radio	540	2.05	0.4223	540	2.62	0.1225	Yes
2-back vs. A-V Radio	540	-9.05	<.0001	540	-6.72	<.0001	Yes

*Direction of differences not consistent

The results were consistent across DRT types for five of the nine comparisons and inconsistent for four of the comparisons, one of which reflected different ordering of conditions. To the extent that 2-back may represent a benchmark level of unacceptable demand, it is noteworthy first that both DRTs revealed 1-back to be different from 2-back and second that the DRTs provided different results for the 2-back versus destination entry comparison. DRT accuracy for destination entry was significantly higher than for 2-back with RDRT but not different for the TDRT. Secondary task by DRT accuracy means, with confidence intervals, are presented in Figure 29.

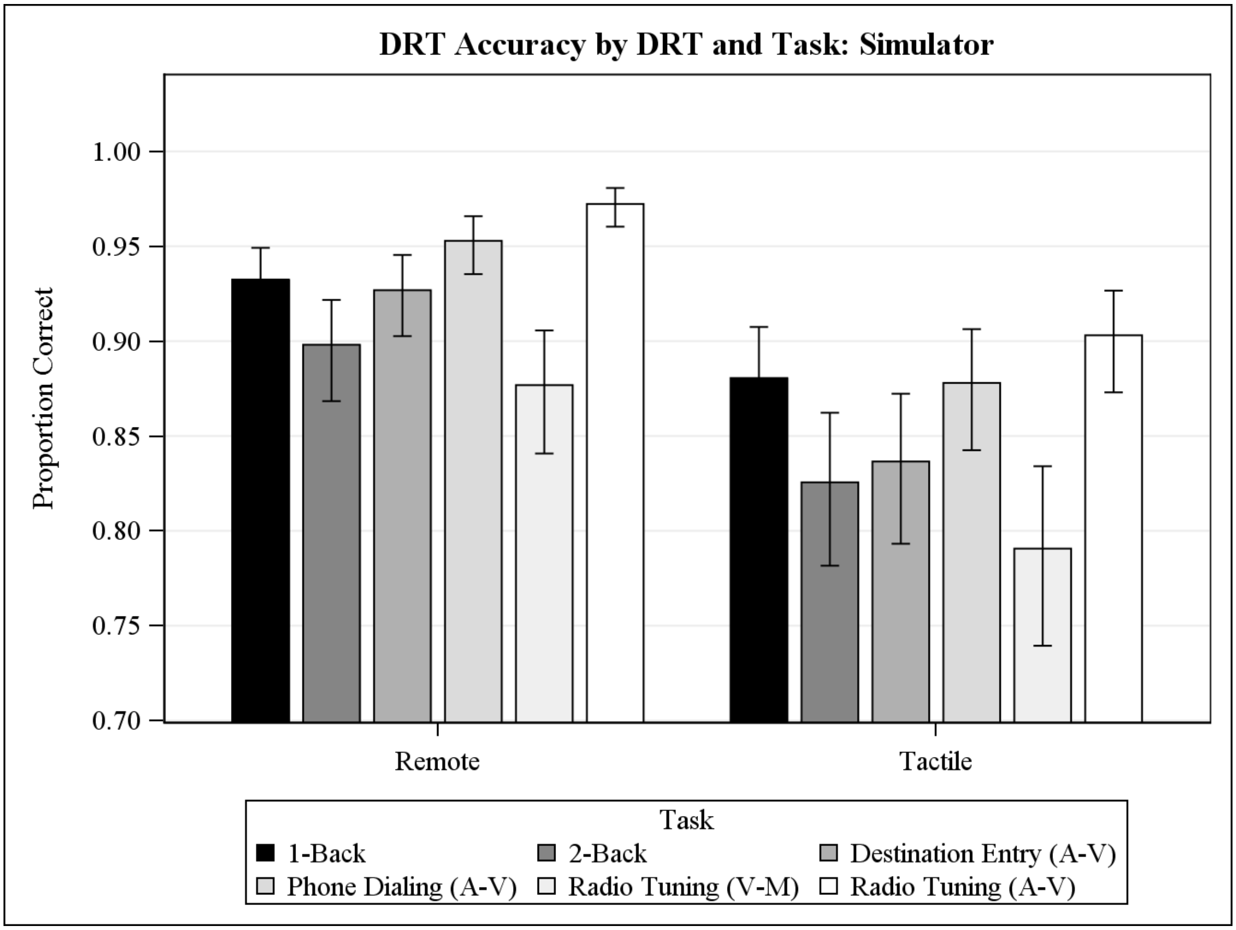


Figure 29. DRT Accuracy by DRT Condition and Task

The overall lower accuracy associated with TDRT responding is apparent in Figure 29 as is the higher variability as reflected in the wider confidence intervals associated with the estimated mean values.

3.1.2.3 Non-Driving Venue DRT RT

An ANOVA was conducted to examine the effects of independent variables on RT in the non-driving venue. As in the driving simulator venue, RT was modeled using a gamma distribution. Independent variables included DRT, age group, and task. The baseline RT was included as a covariate in the model to account for individual differences in RT. The summary of tests for fixed effects is shown in Table 8.

Table 8. ANOVA Fixed Effects: DRT Non-Driving Response Time

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
DRT	1	87	0.12	0.7348
Age Group	3	87	12.63	<.0001
Baseline RT	1	87	68.90	<.0001
Task	5	455	57.45	<.0001
DRT*Task	5	455	1.15	0.3342
DRT*Age Group	3	87	0.25	0.8616
Age Group*Task	15	455	1.39	0.1461

Significant main effects were found for age group and task condition, as well as for the covariate, baseline RT, which reflects individual differences in the participants' response times. Neither the main effect of DRT nor either of the interaction effects involving DRT was significant, which indicates that RT differences were generally consistent across DRT conditions in the non-driving venue.

Table 9 presents the results of planned comparisons involving the 1-back and 2-back conditions.

Table 9. Pairwise Differences Between Task Conditions by DRT Condition (Non-Driving RT)

	RDRT			TDRT			Consistent
	DF	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	455	-5.16	<.0001	455	-3.02	0.0411	Yes
1-back vs. Destination Entry	455	2.79	0.0771	455	3.54	0.0067	No
1-back vs. Phone Dialing	455	3.87	0.0022	455	4.46	0.0004	Yes
1-back vs. V-M Radio	455	-4.45	0.0004	455	-3.80	0.0025	Yes
1-back vs. A-V Radio	455	2.04	0.4184	455	4.83	<.0001	No
2-back vs. Destination Entry	455	7.95	<.0001	455	6.56	<.0001	Yes
2-back vs. Phone Dialing	455	9.03	<.0001	455	7.48	<.0001	Yes
2-back vs. V-M Radio	455	0.71	0.9991	455	-0.77	0.9983	No*
2-back vs. A-V Radio	455	7.20	<.0001	455	7.86	<.0001	Yes

*Direction of differences not consistent

Results were consistent across DRT conditions for six of the nine planned comparisons. Exceptions included the 1-back versus destination entry comparison and the 1-back versus A-V radio tuning comparison, both of which were statistically significant for the TDRT condition but not for the RDRT condition. The 2-back versus V-M radio tuning comparison was consistent in terms of statistical outcome, but the respective minor differences were in different directions. Estimated mean values together with associated confidence intervals are presented in the Figure 30.

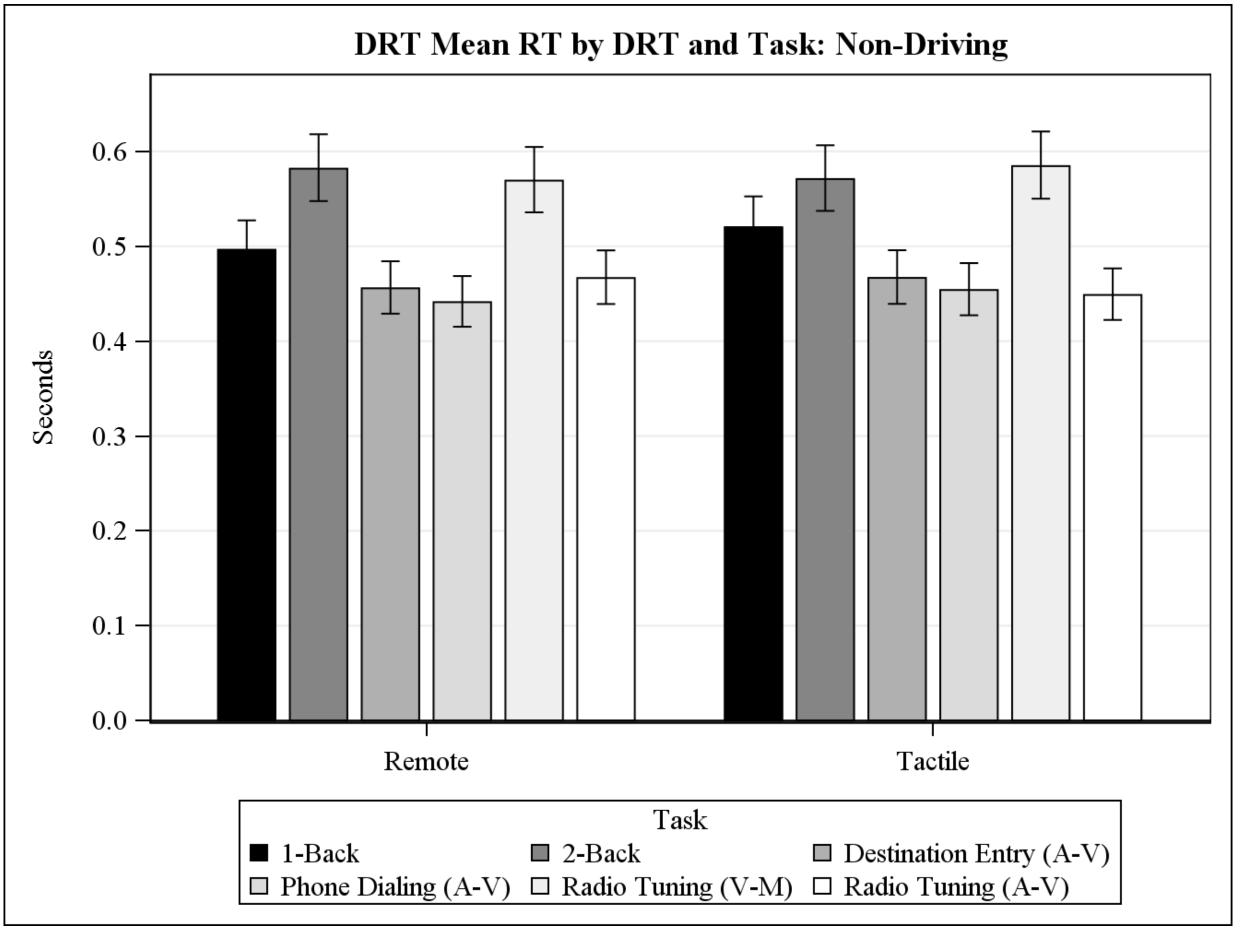


Figure 30. DRT Mean RT by DRT and Task Condition: Non-Driving Venue

Relative to other task conditions, it appears that the participants' responses in the 1-back condition were slower with the TDRT than with the RDRT and that the responses to the 2-back were slightly faster. The slower mean RT values in the 1-back condition resulted in the two discrepant outcomes described above. Differences among other task conditions and 2-back appear greater with RDRT than for TDRT.

3.1.2.4 Non-Driving Venue DRT Accuracy

An ANOVA was conducted to examine the effects of independent variables on DRT accuracy in the non-driving venue. DRT accuracy, defined as the proportion of correct responses, was modeled as a binomial variable. Independent variables included DRT, age group, and task. The baseline Proportion Correct had no effect and was not included in this analysis. The summary of tests for fixed effects is shown in **Error! Reference source not found. Error! Reference source not found.**ANOVA Fixed Effects: DRT Accuracy: Non-Driving Venue

Table 10. ANOVA Fixed Effects: DRT Accuracy: Non-Driving Venue

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
DRT	1	97.92	0.25	0.6184
Age Group	3	100	7.97	<.0001
Task	5	543	46.24	<.0001
DRT*Task	5	543	2.85	0.0150
DRT*Age Group	3	86.61	2.31	0.0821
Age Group*Task	15	543	1.54	0.0866

The DRT main effect was not significant; however, the DRT x Task interaction was significant. Age group and task had significant main effects while the DRT x Age Group and Age Group x Task interactions were not significant. Results of planned comparisons between task conditions and the two benchmark conditions are presented in Table 10.

Table 10. Pairwise Differences Between Task Conditions by DRT Condition (Non-Driving Accuracy)

	RDRT			**TDRT**			**Consistent**
	DF	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	542	2.78	0.0804	542	4.31	0.0003	No
1-back vs. Destination Entry	542	-6.55	<.0001	542	-2.10	0.3918	No
1-back vs. Phone Dialing	542	-6.29	<.0001	542	-3.79	0.0026	Yes
1-back vs. V-M Radio	542	-3.34	0.0140	542	-1.70	0.6933	No
1-back vs. A-V Radio	542	-6.53	<.0001	542	-3.02	0.0409	Yes
2-back vs. Destination Entry	542	-8.41	<.0001	542	-5.60	<.0001	Yes
2-back vs. Phone Dialing	542	-8.15	<.0001	542	-6.84	<.0001	Yes
2-back vs. V-M Radio	542	-5.66	<.0001	542	-5.31	<.0001	Yes
2-back vs. A-V Radio	542	-8.33	<.0001	542	-6.27	<.0001	Yes

Results were consistent for six of the nine planned comparisons. The difference between 1-back and 2-back conditions was significant in the TDRT condition but not in the RDRT condition. The 1-back versus destination entry comparison had different outcomes by DRT condition, as did the 1-back versus V-M radio tuning comparison. Mean values and confidence intervals are presented in Figure 31.

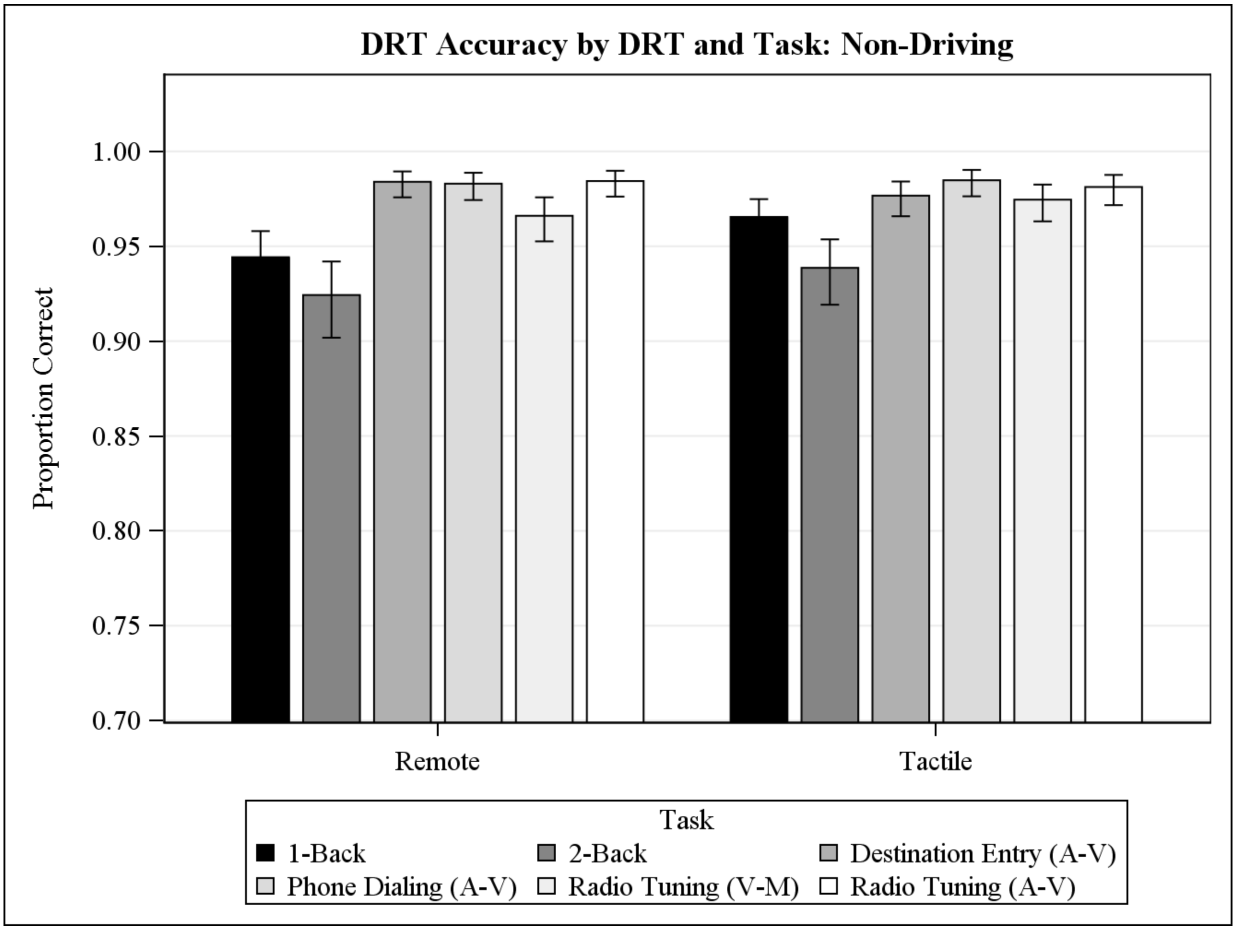


Figure 31. DRT Mean Accuracy by DRT and Task Condition: Non-Driving Venue

Two differences are apparent; first, the 1-back and 2-back accuracy values were lower for the RDRT than for the TDRT; and second, the V-M radio tuning task accuracy was lower in RDRT relative to TDRT.

3.1.3 Driving Simulator versus Non-Driving Test Venue

3.1.3.1 Response Time

Data from both venues were combined to examine differences between venues. An ANOVA was performed on the combined data, which comprised mean RTs and accuracy proportions for each task condition. For the first analysis, RT was the dependent variable; independent variables included DRT, age group, task, and venue. Interactions addressing specific questions were included in the analysis model. Results of tests for fixed effects are presented in

Table 11.

Table 11. Summary of Fixed Effects for Data Combined From Both Venues: Response Time

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
DRT	1	178	18.00	<.0001
Age Group	3	178	15.98	<.0001
Baseline RT	1	178.1	90.61	<.0001
Task	5	923.1	74.67	<.0001
Venue	1	178.2	14.33	0.0002
DRT*Task	5	923.1	1.90	0.0914
DRT*Age Group	3	178	0.20	0.8963
DRT*Venue	1	178	12.27	0.0006
Age Group*Venue	3	178	0.82	0.4829
Task*Venue	5	923.1	9.96	<.0001
Age Group*Task	15	923.1	2.43	0.0017
DRT*Task*Venue	5	923.1	1.30	0.2635

As seen in previous analyses, DRT, age group, and task had significant main effects. Of interest were effects associated with venue, which included the significant main effect of venue plus significant interaction effects associated with DRT x Venue and Task x Venue. Means and associated confidence intervals for the RT values associated with combinations of DRT and test venue are presented in Figure 32. They reveal the relatively large differences between DRT types apparent in the simulator that were not apparent in the non-driving venue. The tactile DRT stimuli were apparently not detected as readily in the simulator as in the non-driving venue perhaps because participants assigned a lower priority to responding to a stimulus that did not have an established connection with safe driving. Signal strength was not likely an explanation because the RT means in the baseline condition were comparable across DRT types in the simulator.

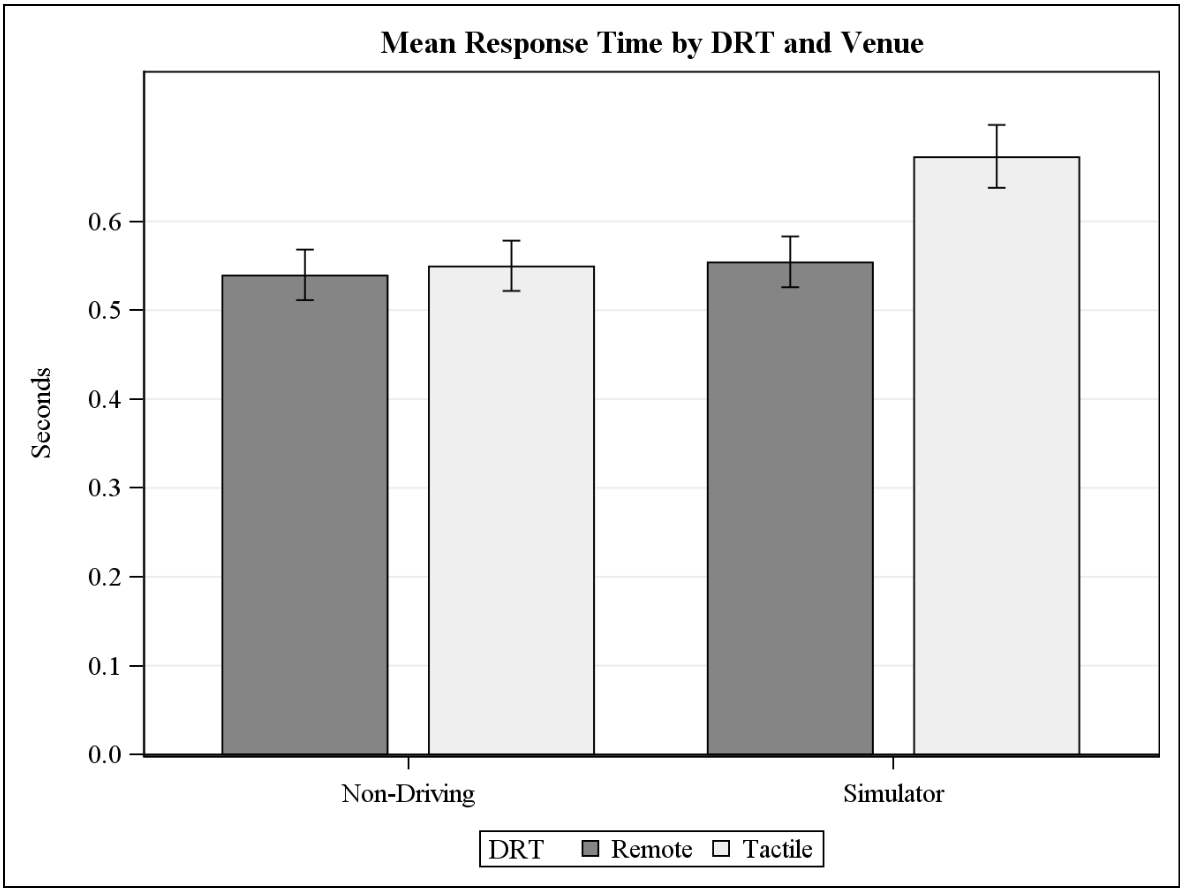


Figure 32. DRT Mean RT Means by Venue and DRT Condition

The results of planned pairwise comparisons for the venue and task interaction are presented in Table 12. Comparisons were selected using the two benchmark task conditions 1-back and 2-back. Specifically, comparisons were selected to see if differences involving comparisons with one of the two benchmark conditions differed between test venues.

Table 12. Results of Planned Comparisons for Selected Combinations of Venue and Task (Response Time)

	Non-Driving			**Simulator**			**Consistent**
	DF	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	923	-5.97	<.0001	923	-3.64	0.0052	Yes
1-back vs. Destination Entry	923	4.58	<.0001	923	-1.46	0.8355	No*
1-back vs. Phone Dialing	923	6.04	<.0001	923	1.01	0.9852	No
1-back vs. V-M Radio	923	-6.03	<.0001	923.5	-7.66	<.0001	Yes
1-back vs. A-V Radio	923	4.97	<.0001	814.4	-21.71	<.0001	No*
2-back vs. Destination Entry	923	10.55	<.0001	923	2.18	0.3301	No
2-back vs. Phone Dialing	923	12.01	<.0001	923	4.65	<.0001	Yes
2-back vs. V-M Radio	923	-0.06	1.0000	923.5	-4.04	0.0009	Yes
2-back vs. A-V Radio	923	10.94	<.0001	923	5.19	<.0001	Yes

*Direction of differences not consistent

Of the nine selected comparisons, five had consistent outcomes across the two venues and four had different outcomes. Comparisons with different outcomes in different venues included: 1-back versus destination entry; 1-back versus phone dialing; 1-back versus A-V radio tuning; and 2-back versus destination entry. Two of the comparisons had differences in different directions, reflecting the finding that three of the in-vehicle tasks had much faster DRT response times in the non-driving (ND) venue than in the simulator (Sim). Means for this interaction effect are presented in Figure 33.

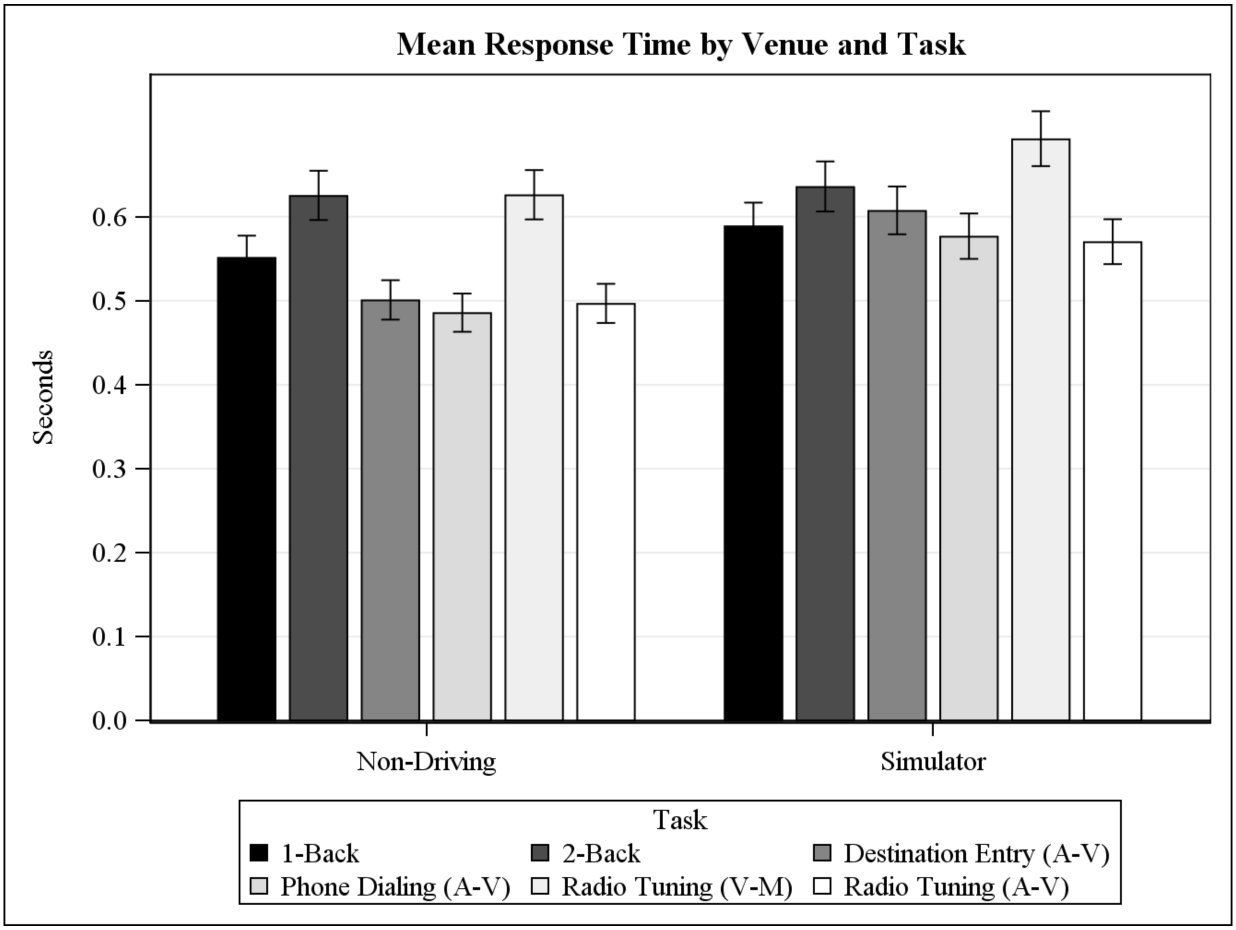


Figure 33. Mean Response Time by Test Venue and Task Condition

The DRT x Task x Venue interaction was not significant. However, it was apparent that the mean RT values associated with the simulator task conditions were slower than those in the non-driving venue, due to the consistently slower responses observed in the simulator using the tactile DRT. Because of the differences between DRT conditions in the simulator, these data were separated further by DRT condition, as shown in Figure 34.

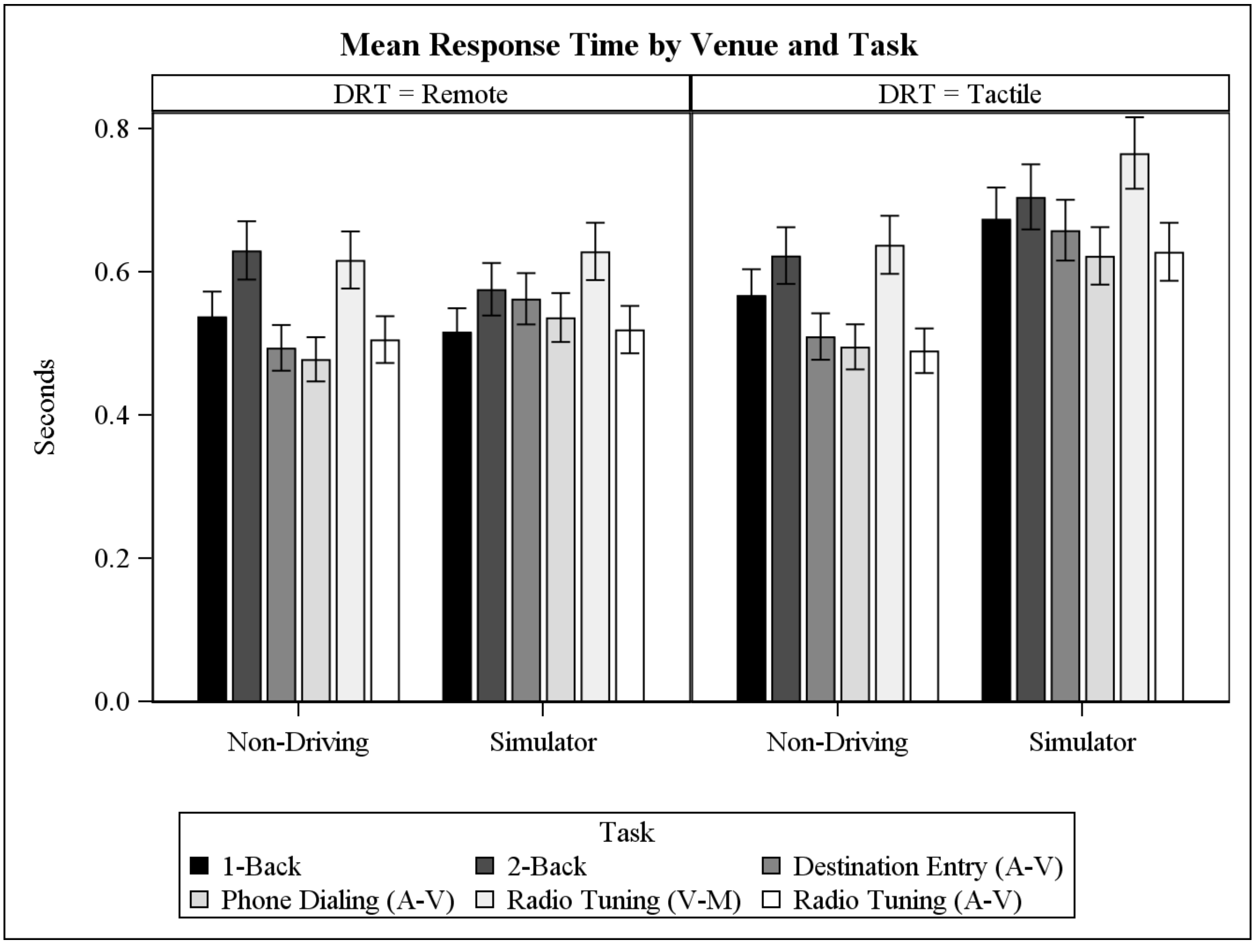


Figure 34. Mean Response Time by Venue, Task, and DRT Condition

The generally slower responding in the simulator TDRT condition is apparent in this figure (Figure 34). Overall, the differences among task conditions were larger in the non-driving venue than in the simulator venue. Thus, the requirement to perform a driving task serves to minimize the differences among in-vehicle tasks that were apparent when no driving task was required. This may reflect the smaller pool of attentional resources presumably available in the driving simulator for in-vehicle task performance, given the demands of the driving task, relative to those in the non-driving venue, which has no such demands.

Considering the RDRT means on the left side of Figure 34, the most apparent difference between test venues is that the differences between conditions appear consistently greater in the ND venue. The difference between 1-back and 2-back conditions is largest with RDRT in the ND venue. However, the size of the differences is not the only difference between test venues with RDRT. The ordering of tasks differs between test venues. In the ND venue, the destination entry and phone dialing task mean RT values were lowest, suggesting that these were the least demanding tasks, followed by A-V radio tuning and then 1-back. However, with the RDRT in the simulator, A-V radio tuning and 1-back had the fastest RT mean values, followed by phone dialing, destination entry, and 2-back. If faster RT means reflect better performance, it appears that with the additional driving task demands in the simulator, participants performed better on the 1-back and 2-back tasks than in the ND venue. In contrast, simulator participants were slower responding to DRT targets for the voice-based in-vehicle tasks (destination entry, phone dialing, A-V radio tuning) than in the ND venue. The different ordering of tasks in different test venues may reflect differences between performing tasks intermittently, as would be required in the simulator, versus continuously, as would be allowed in the ND venue. A similar pattern of task ordering differences appears to exist between venues for the TDRT.

This pattern was most notable for the destination entry, phone dialing, and A-V radio tuning tasks, which were associated with shorter RTs relative to the other conditions in the non-driving venue than in the simulator venue. Of interest is the difference between the four in-vehicle task conditions and the two benchmark conditions, which are considered surrogates for acceptable (i.e., 1-back) and unacceptable (i.e., 2-back) levels of attentional demand. In the non-driving venue, the mean RT values for destination entry, phone dialing, and A-V radio tuning tasks were all faster than both benchmarks, which could be taken to mean that these tasks are less demanding than either of the benchmarks. In the simulator, the differences between these tasks and the benchmark tasks are much smaller. Consider for example the destination entry task, which has RT values that are significantly faster than 2-back in the non-driving venue, but statistically not different from 2-back in the driving simulator venue. If 2-back were to be considered the threshold of unacceptable attentional demand, the outcomes for this task would differ according to venue.

3.1.3.2 Accuracy

Distributions of DRT accuracy reveal two differences between venues. First, the accuracy distributions reveal consistently higher levels of variability in the simulator than in the non-driving venue. This undoubtedly reflects the fundamental differences between the venues. Second, the accuracy scores appear to have been constrained by the ceiling associated with perfect performance in the non-driving venue. In addition, the accuracy distributions revealed differences between DRT types in the simulator venue.

To examine these differences more formally, an ANOVA was performed on the combined data, with DRT accuracy as the dependent variable. Independent variables included DRT, age group, task, and venue. Baseline accuracy was included as a covariate in this analysis. Interactions addressing specific questions were included in the analysis model. Results of tests for fixed effects are presented in Table 13.

Table 13. Summary of Fixed Effects for Data Combined From Both Venues: Accuracy

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
DRT	1	185	8.51	0.0040
Age Group	3	179.8	20.36	<.0001
Baseline Correct	1	151.6	3.95	0.0486
Task	5	1101	71.17	<.0001
Venue	1	185	89.06	<.0001
DRT*Task	5	1101	5.51	<.0001
DRT*Age Group	3	170.3	0.68	0.5647
Age Group*Task	15	1101	7.12	<.0001
DRT*Venue	1	183.7	10.94	0.0011
Age Group*Venue	3	173.4	0.19	0.9040
Task*Venue	5	1101	30.71	<.0001
DRT*Task*Venue	5	1101	1.05	0.3871

The results indicated significant main effects of DRT, age group, task, and venue. The following interaction effects were also significant: DRT x Task, Age Group x Task, DRT x Venue, and Task x Venue. Estimates of mean DRT Accuracy values with confidence intervals associated with the DRT x Venue interaction effect are presented in Figure 35.

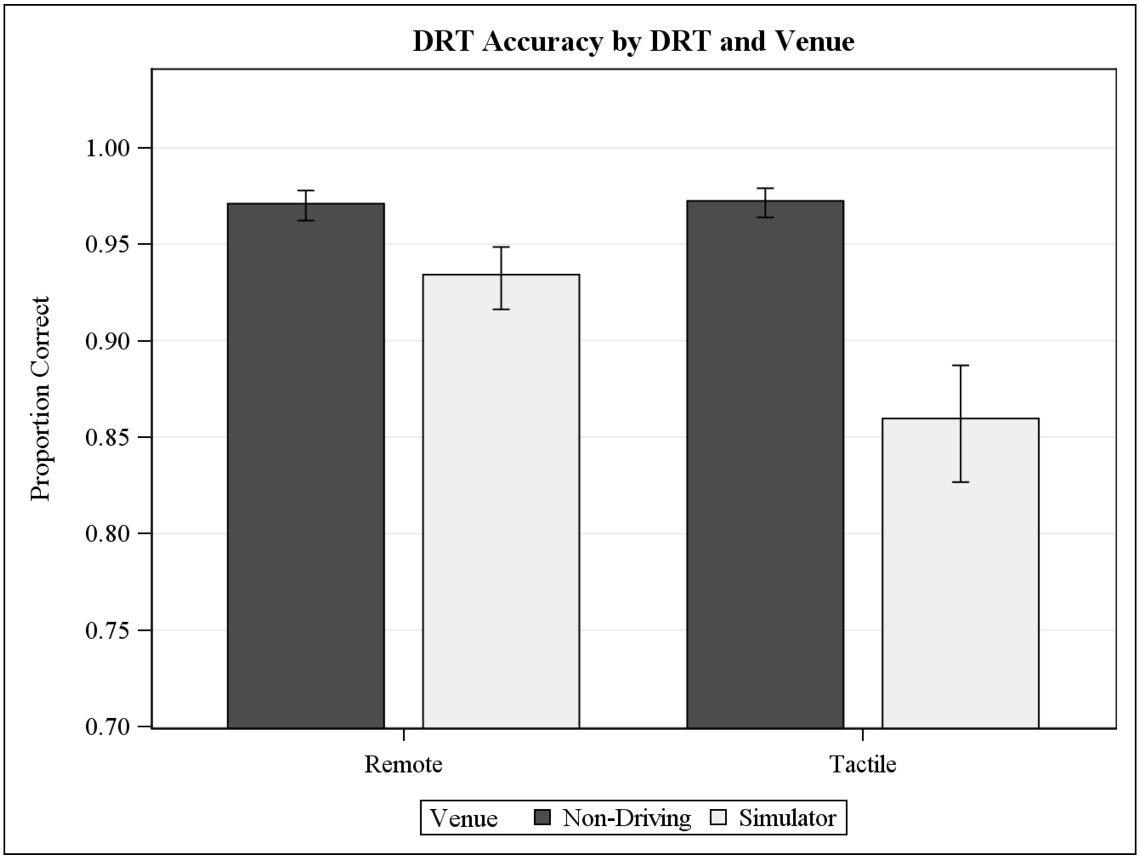


Figure 35. Mean DRT Accuracy by Venue and DRT Condition

The mean overall accuracy was approximately 97 percent in the non-driving venue for both DRTs; however, the simulator venue overall accuracy was lower in both DRT conditions with a much bigger decrease associated with the TDRT (86% for TDRT vs. 93% for RDRT in the Sim).

Estimated mean DRT accuracy values associated with the DRT x Task interaction effect are presented in Figure 36.

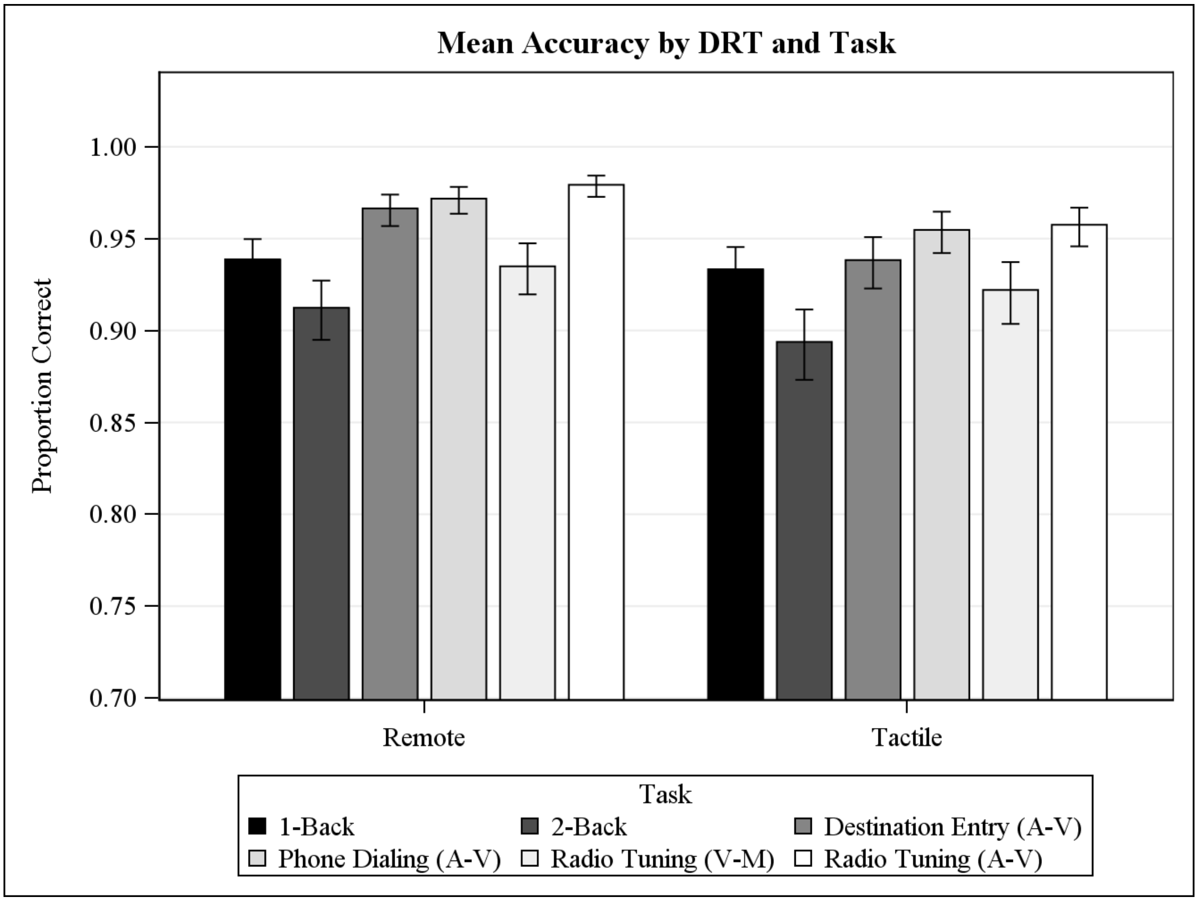


Figure 36. Mean Accuracy by Task and DRT Condition (Both Venues)

Most apparent from this interaction effect are differences in the respective differences among task conditions with different DRTs. Overall accuracy is slightly higher with RDRT for all task conditions. Mean DRT accuracy associated with the significant Venue x Task interaction effect is presented in Figure 37.

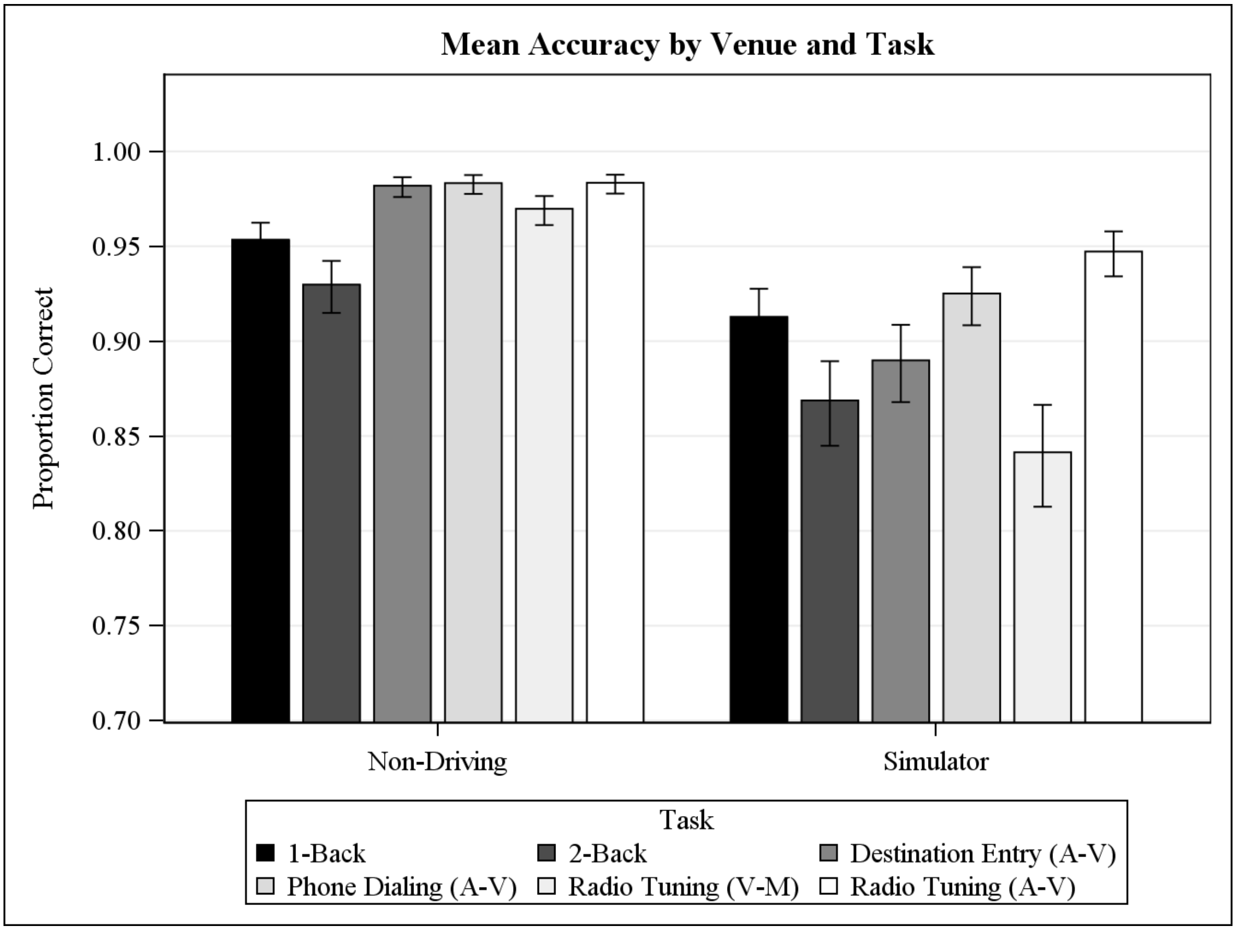


Figure 37. Mean Accuracy by Venue and Task Condition (Both Venues)

Figure 37 shows how the consistently high accuracy associated with the ND venue reduces sensitivity for detecting differences between task conditions. The lower accuracies in the simulator allow for differences between task conditions to emerge.

Results of planned comparisons associated with the Task x Venue interaction are presented in Table 14. These comparisons examine the consistency of results across the two venues.

Table 14. Results of Planned Comparisons for Selected Combinations of Venue and Task (Accuracy)

	ND			Sim			Consistent
	DF	**t Value**	**Adjusted P**	**DF**	**t Value**	**Adjusted P**	
1-back vs. 2-back	1101	4.89	<.0001	1101	6.68	<.0001	Yes
1-back vs. Destination Entry	1101	-7.21	<.0001	1101	3.36	0.0136	No*
1-back vs. Phone Dialing	1101	-7.57	<.0001	1101	-1.93	0.5233	No
1-back vs. V-M Radio	1101	-3.81	0.0026	1101	9.30	<.0001	No*
1-back vs. A-V Radio	1101	-7.54	<.0001	1101	19.64	<.0001	No*
2-back vs. Destination Entry	1101	-10.85	<.0001	1101	-2.76	0.0843	No
2-back vs. Phone Dialing	1101	-11.10	<.0001	1101	-7.66	<.0001	Yes
2-back vs. V-M Radio	1101	-7.94	<.0001	1101	3.25	0.0193	No*
2-back vs. A-V Radio	1101	-11.00	<.0001	1101	-10.85	<.0001	Yes

*Direction of differences not consistent

The assessments of consistency were based on two criteria; first, whether the differences were in the same direction, which is reflected in the signs associated with the t values and second, whether the differences had the same statistical outcome. Results were consistent across venues for three of the nine selected pairwise comparisons. Four of the comparisons were inconsistent because the differences were not in the same direction.

3.1.4 Age Group Effects

Age effects were examined to determine if drivers of different ages responded differently to different tasks in different test venues or by using different DRT devices. The analyses reported above included main and interaction effects of Age and the results were generally consistent across the various analyses:

1. Age main effects were present in each analysis, reflecting the expected finding that RT increases and accuracy decreases with increasing age, generally.
2. Age x Task interactions were found in both simulator metrics but not in the non-driving metrics. These effects suggest that the ordering of task means differ for the different age groups, primarily in the simulator.

These significant effects will be examined here. First, the distributions of the DRT summary metrics by age group and test venue are presented. Figure 38 shows the DRT response time distributions for the simulator venue, while Figure 39 shows the non-driving venue distributions.

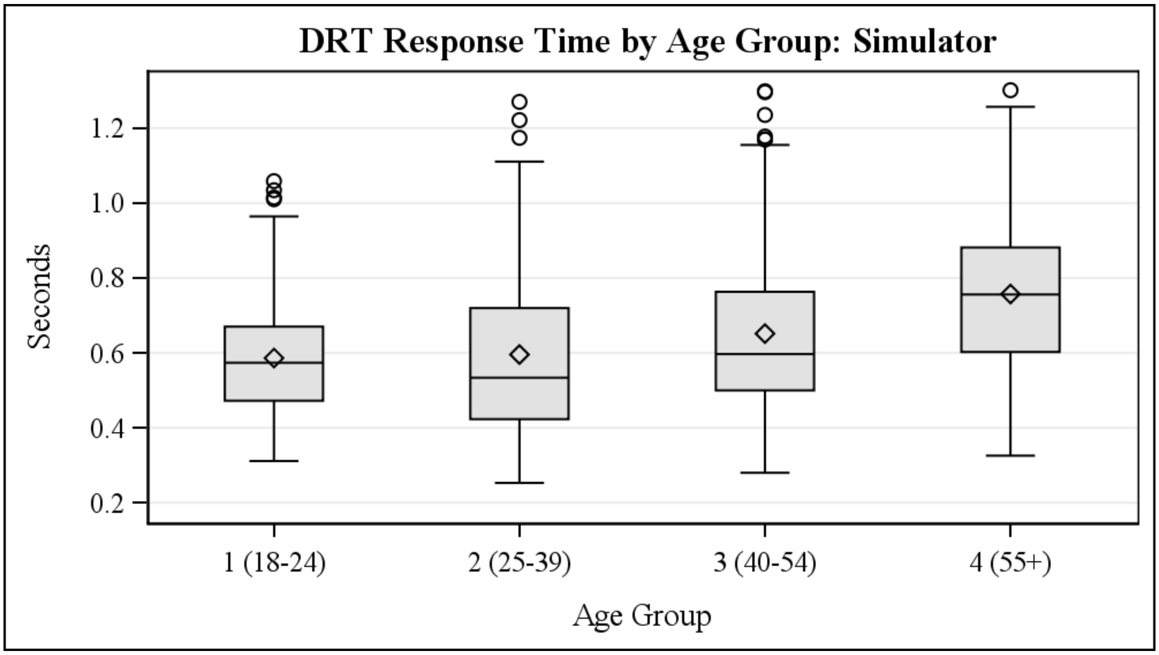


Figure 38. DRT Mean Response Times by Age Group (Simulator)

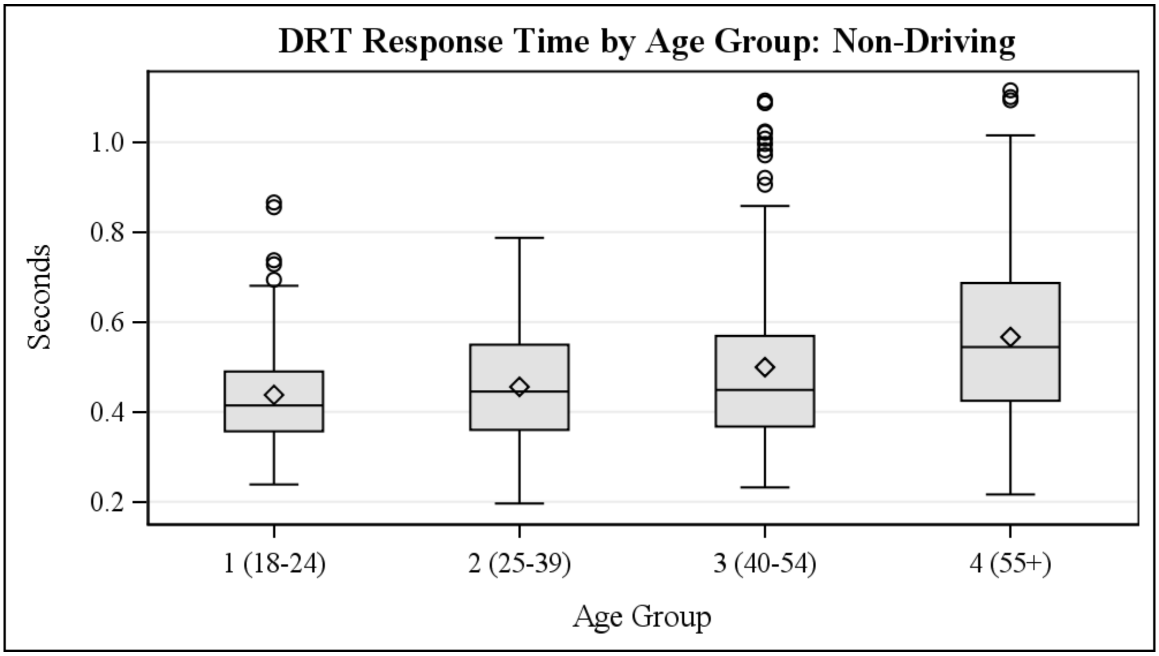


Figure 39. DRT Mean Response Times by Age Group (Non-Driving)

Data from both venues reveal the expected increasing mean RT values with increasing age. Median values remained relatively constant across the first three age groups, increasing in the fourth group (i.e., 55+); however, mean values were pulled above the median by outliers in most groups. The 40-54 non-driving group had the most outliers. Also evident is the increasing variability in the interquartile values (box sizes) among participants with increasing age. Figure 40 shows the DRT accuracy distributions for the simulator venue, while Figure 41 shows the non-driving venue distributions.

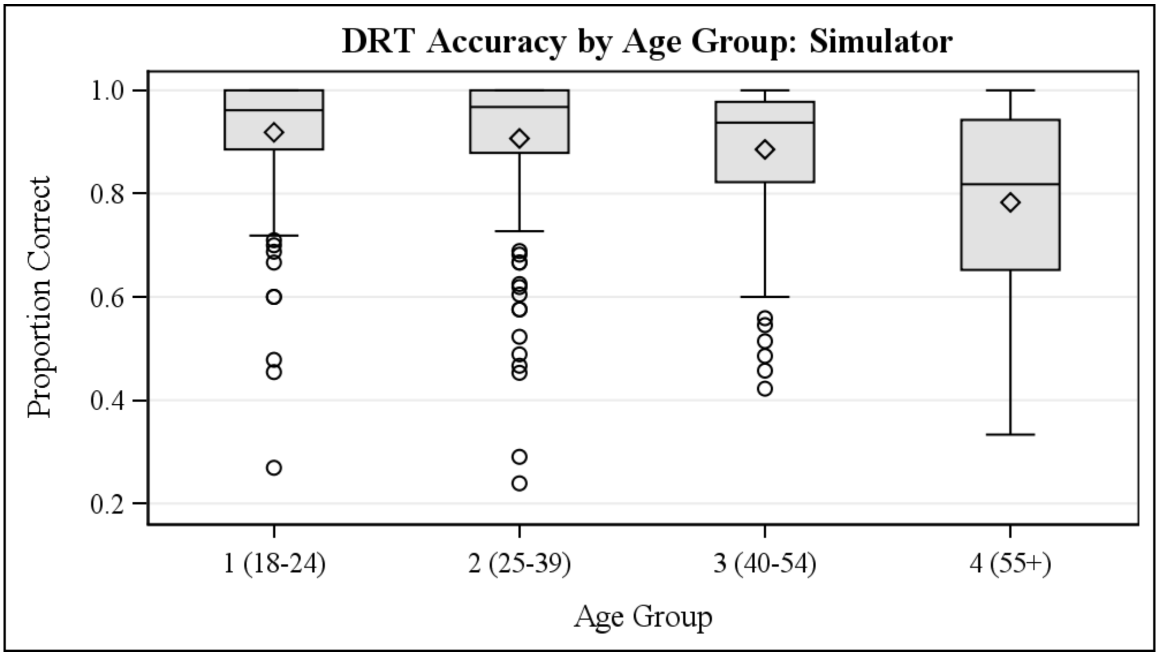


Figure 40. DRT Detection Accuracy by Age Group (Simulator)

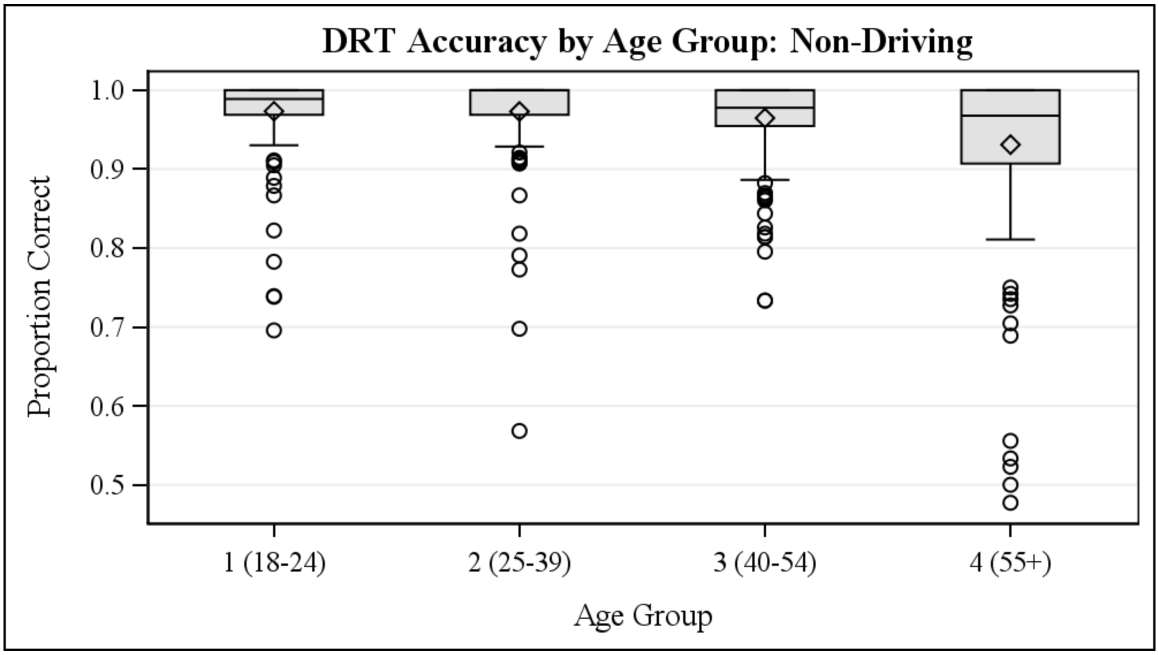


Figure 41. DRT Detection Accuracy by Age Group (Non-Driving)

Detection accuracy decreased with increasing age in both venues, but was more apparent in the simulator data, reflecting the increased difficulty of target detection while driving. Detection accuracy was generally more variable in the simulator as reflected by the larger boxes. Extremely low detection accuracy values suggest poor overall DRT performance; however, the ISO protocol does not have a minimum accuracy threshold for accepting response time data.

The DRT x Age interaction effects were generally not statistically significant; however, as shown in the following two figures (Figure 42 response time and Figure 43 accuracy), the TDRT was associated with more pronounced age effects than the RDRT in the simulator venue. TDRT also had more outliers associated with poor accuracy than the RDRT. This pattern of differences between DRTs was not apparent in the non-driving venue (data not shown).

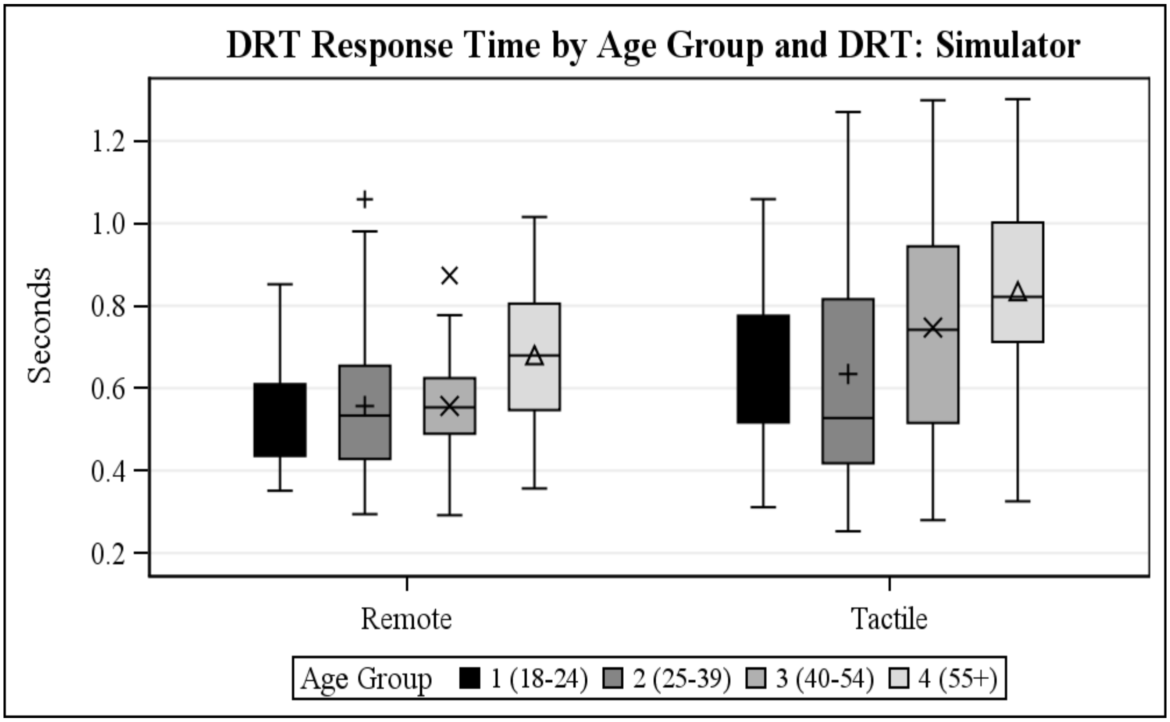


Figure 42. Distributions of DRT Response Time by Age Group and DRT Condition (Simulator)

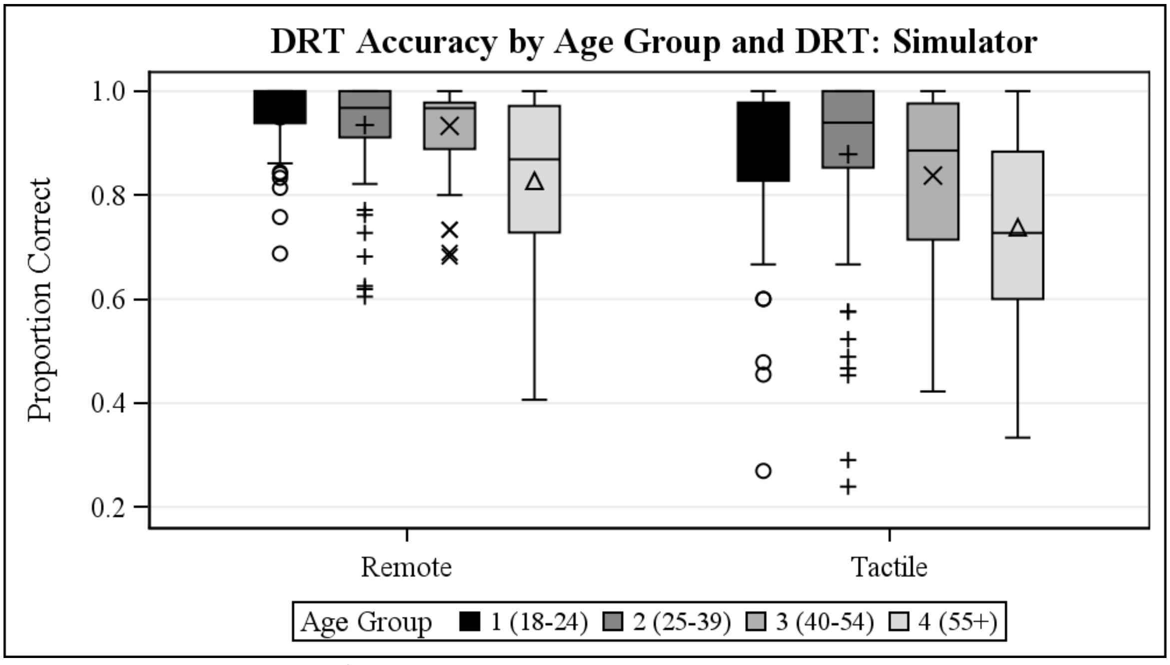


Figure 43. Distributions of DRT Accuracy by Age Group and DRT Condition (Simulator)

Figure 44 presents means associated with the significant age group x Task interaction for simulator DRT RTs.

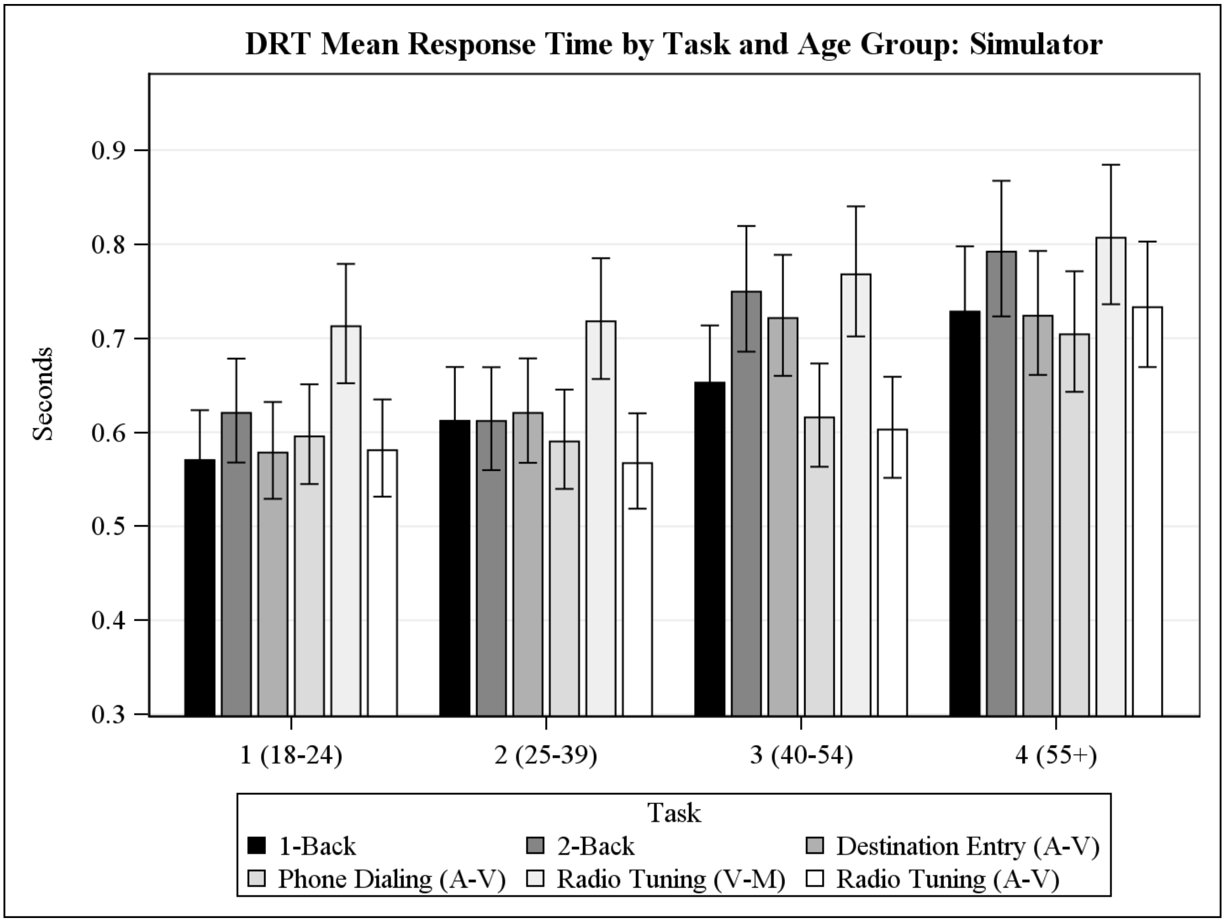


Figure 44. Mean RT by Age Group and Task Condition (Simulator)

The progressive effects of age on response time are evident; however, there are some differences between age groups. Consider the 1-back and 2-back tasks, which have been the focus of analyses due to their use as benchmark tasks. The pattern of differences between these two task conditions differs among the four age groups. The difference in mean RT values is greatest for the 40-54 group, followed by the 55+ group. The same pattern is apparent for the 18-24 group. However, for the 25-39 group, the means are close to identical for these two conditions. Differences between mean RT values for the destination entry and phone dialing tasks also differ among age groups with a large difference evident in the 40-54 group that is not apparent elsewhere.

The corresponding means associated with the significant Age Group x Task interaction effect for DRT accuracy in the simulator venue are presented in Figure 45.

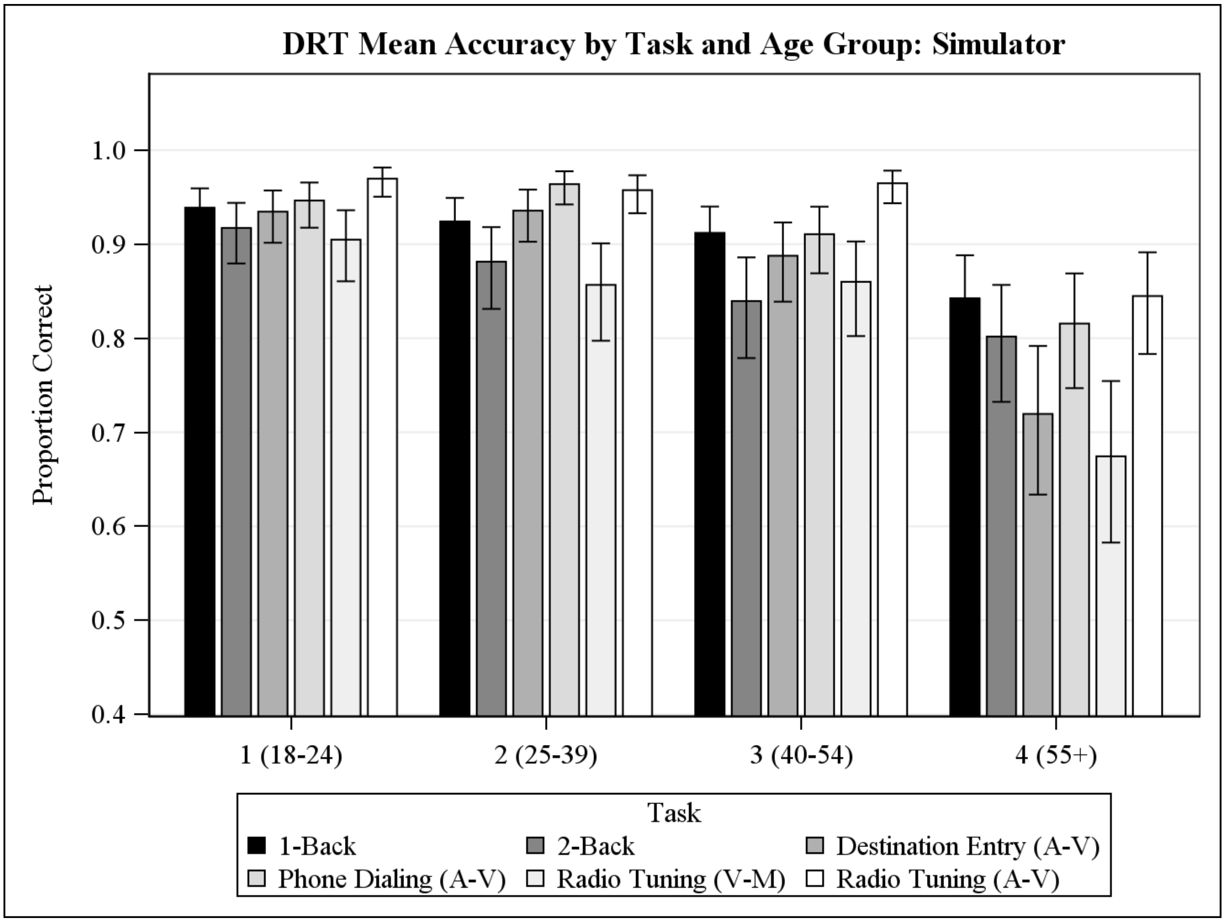


Figure 45. Mean Accuracy by Age Group and Task Condition (Simulator)

DRT detection accuracy decreased with increasing age; however, the changes between age groups differed among the task conditions. Generally, the 18-24-year-old participants could maintain relatively higher and more consistent accuracy levels than participants in the other age groups. More specifically, DRT accuracy associated with the A-V radio tuning task remained high for the first three groups, decreasing only in the fourth group. Accuracy associated with the 1-back task followed a similar pattern. Accuracy associated with other tasks (e.g., 2-back, V-M radio tuning) decreased more gradually across the 4 groups. Differences in accuracy among the task conditions were greatest for the oldest group. For example, the difference between the V-M radio tuning and the A-V radio tuning task accuracy means was greatest for the older groups. Ordering of task accuracies also differed among age groups. The difference between 2-back and destination entry conditions was reversed for the older (i.e., 55+) drivers, relative to the other three age groups. Note also the difference between 2-back and V-M radio tuning conditions, which differed considerably for the older (i.e., 55+) drivers but not for the younger three groups. To the extent that DRT accuracy may reflect task demand, this pattern suggests that participants over 55 found the destination entry and V-M radio tuning tasks to be more demanding than the 2-back benchmark task.

3.1.5 Consistency of Results Across Guidelines Groups

3.1.5.1 Simulator Response Time

ANOVAs were computed to examine differences in DRT performance across Distraction Guidelines groups. The Glimmix procedure (SAS, V9.4), with analysis models described above, was used for this purpose. Task and Guidelines Group were the independent variables. Table 15 presents the results of tests of fixed effects for DRT RT in the simulator.

Table 15. Summary of Fixed Effect Tests (Simulator RT)

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
Task	6	478	118.36	<.0001
Guidelines Group	3	79.99	5.54	0.0017
Guidelines Group*Task	18	478	2.65	0.0003

The results indicated significant main effects of task, GL Group and the GL Group x Task interaction effect. The main effect of GL Group is due to the differences between DRT types; post hoc tests showed no differences between groups with the same DRT. Means and confidence intervals associated with the GL Group x Task interaction are presented in Figure 46.

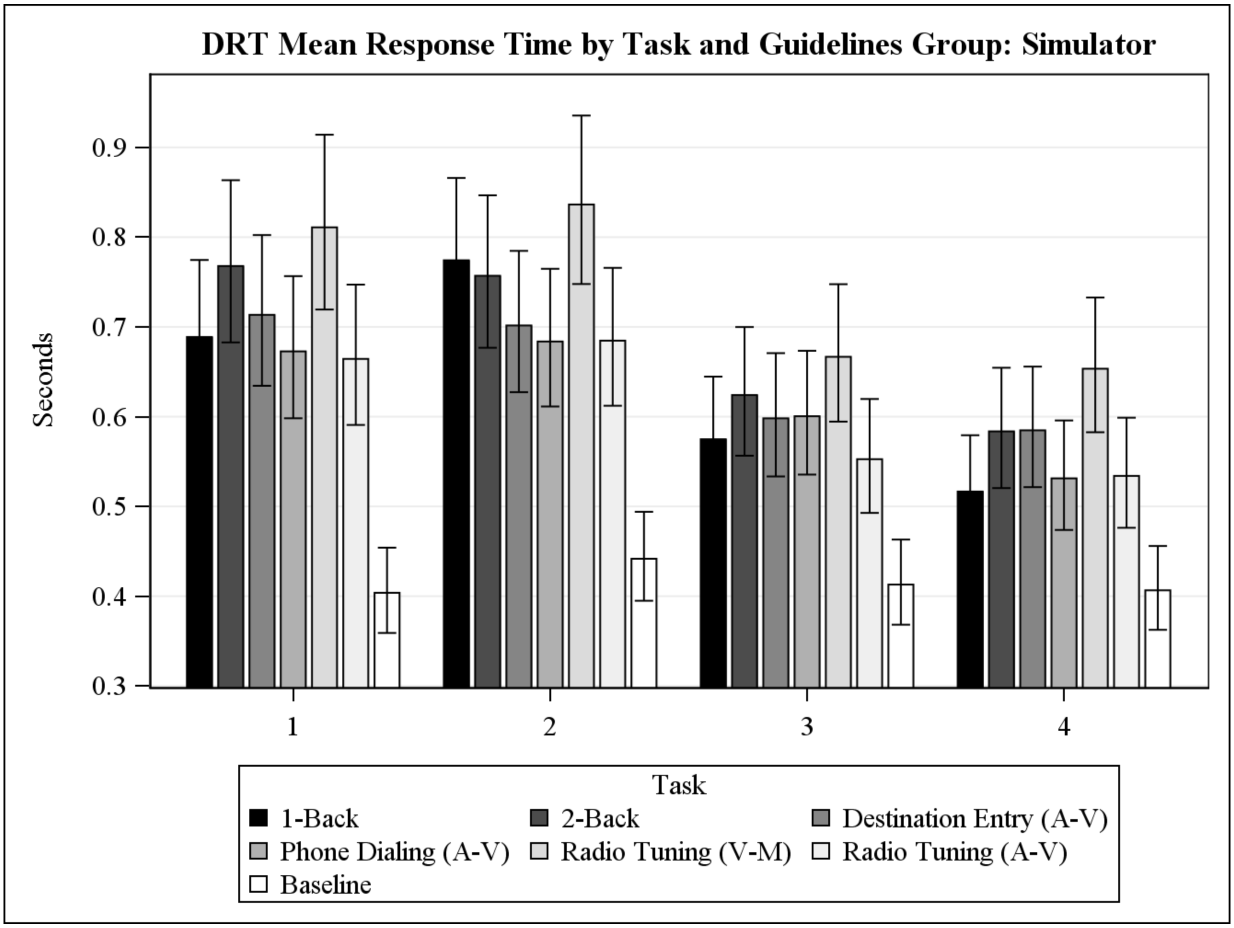


Figure 46. DRT Mean RT by Task and GL Group: Simulator

The main differences apparent in Figure 46 are between the first two and last two groups, which reflect differences due to DRT types. Comparisons of mean RT values reflecting ordering of tasks within DRT types (Groups 1 vs. 2) and (Groups 3 vs. 4) are of interest. The main difference between Groups 1 and 2 is the different ordering of 1-back and 2-back means. Group 1 reveals the predicted effect, with longer RTs for 2-back relative to 1-back; however, the corresponding difference in Group 2 is very small and in the opposite direction. With respect to Groups 3 and 4, it is apparent that the difference between destination entry and phone dialing task means differs between these groups. Destination entry had longer mean RT than the phone dialing task in Group 4 while these corresponding means were essentially equivalent in Group 3.

3.1.5.2 Simulator Accuracy Data

Results for fixed effects associated with the ANOVA computed for DRT detection accuracy in the simulator are presented in Table 16. Means associated with the GL Group x Task interaction effect are shown in Figure 47.

Table 16. Summary of Fixed Effect Tests (Simulator Accuracy)

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
Task	6	558	71.39	<.0001
Guidelines Group	3	83.28	4.01	0.0102
Guidelines Group*Task	18	558	3.70	<.0001

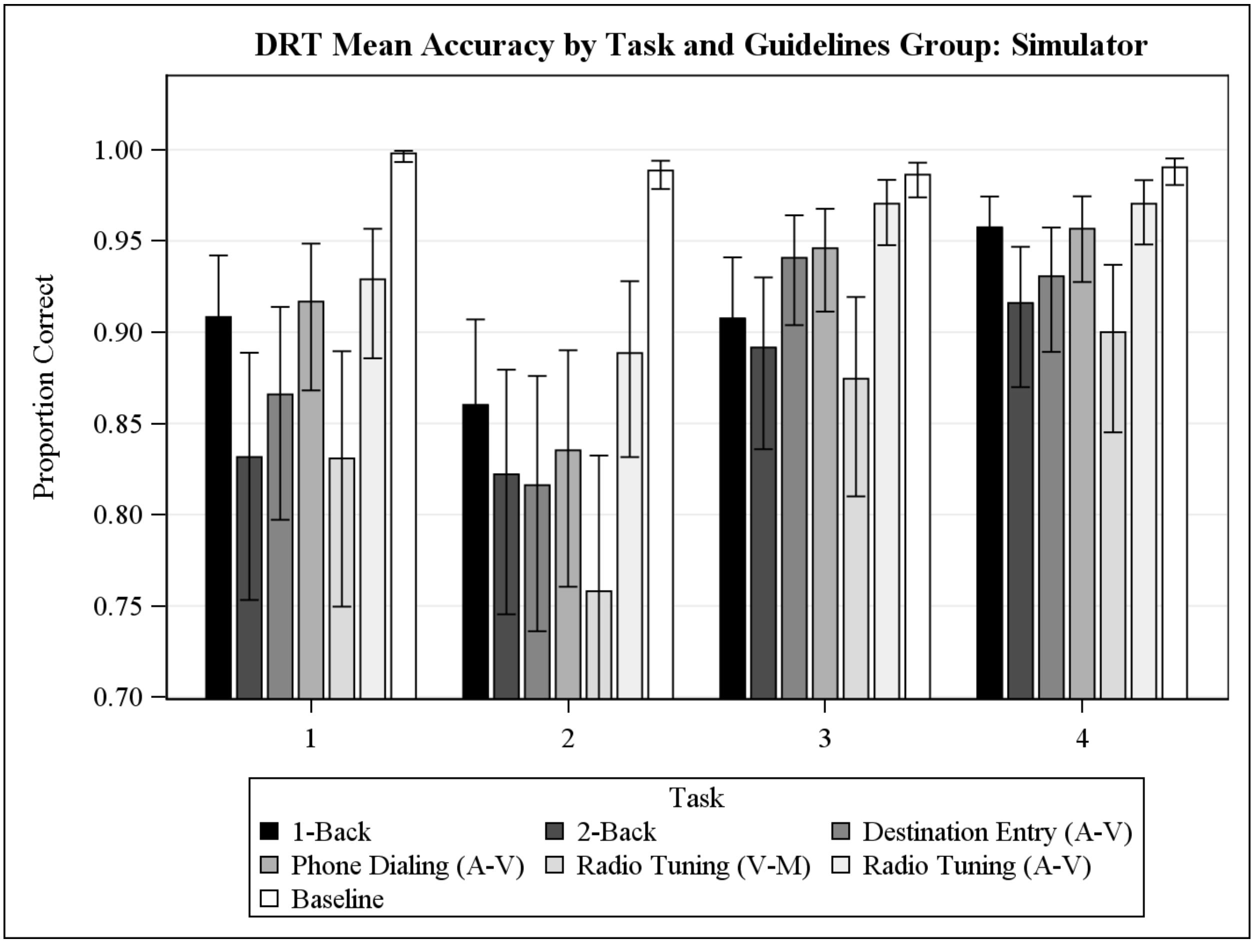


Figure 47. DRT Mean Accuracy by Task and GL Group: Simulator

Differences among Guidelines Groups again reflect differences between the DRT conditions, which were Tactile for Groups 1 and 2 and Remote (i.e., Visual) for Groups 3 and 4. Within the Tactile Groups, the general overall detection accuracy for Group 2 appeared consistently lower than for Group 1, which may reflect differences among individuals in the respective groups. The magnitude of the difference between 1-back and 2-back conditions differs across all groups.

3.1.5.3 Non-Driving Response Time

A summary of tests of fixed effects for Non-Driving DRT RT means is presented in the following table (Table 17). The results indicate no differences among Guidelines Groups or no interaction between tasks and Guidelines groups. For this metric in the non-driving venue, results were consistent across all four GL groups. The absence of a GL Group x Task interaction suggests that the ordering of tasks was relatively consistent within each GL group. Tests were not performed to assess whether the results of statistical tests associated with differences between pairs of tasks within each of the groups were consistent with the results obtained with the aggregated dataset presented earlier. Such comparisons would address questions relating to the effects of different sample sizes (*N* = 96 vs. *N* = 24) on test outcomes.

Table 17. Summary of Fixed Effect Tests (Non-Driving RT)

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
Task	6	552	175.49	<.0001
Guidelines Group	3	92	0.49	0.6896
Guidelines Group*Task	18	552	0.88	0.6048

Figure 48 presents the mean DRT RT values by GL group and task for the non-driving venue. The absence of the GL Group x Task interaction effect indicates that differences between GL groups are minor and with minor exceptions the ordering of task condition means and the apparent differences among task conditions appear generally consistent across GL Groups. For example, differences between 1-back and 2-back condition means are in the same direction across groups while differing slightly in magnitude. Differences in mean RT values associated with destination entry and phone dialing conditions are not consistent across groups. Small differences are apparent for GL Groups 1 and 3, but not for Groups 2 and 4. Although statistical testing was not done to formally compare these differences, the absence of the interaction effect and the relatively larger confidence interval boundaries associated with means for small samples suggest that none of these differences would be statistically significant.

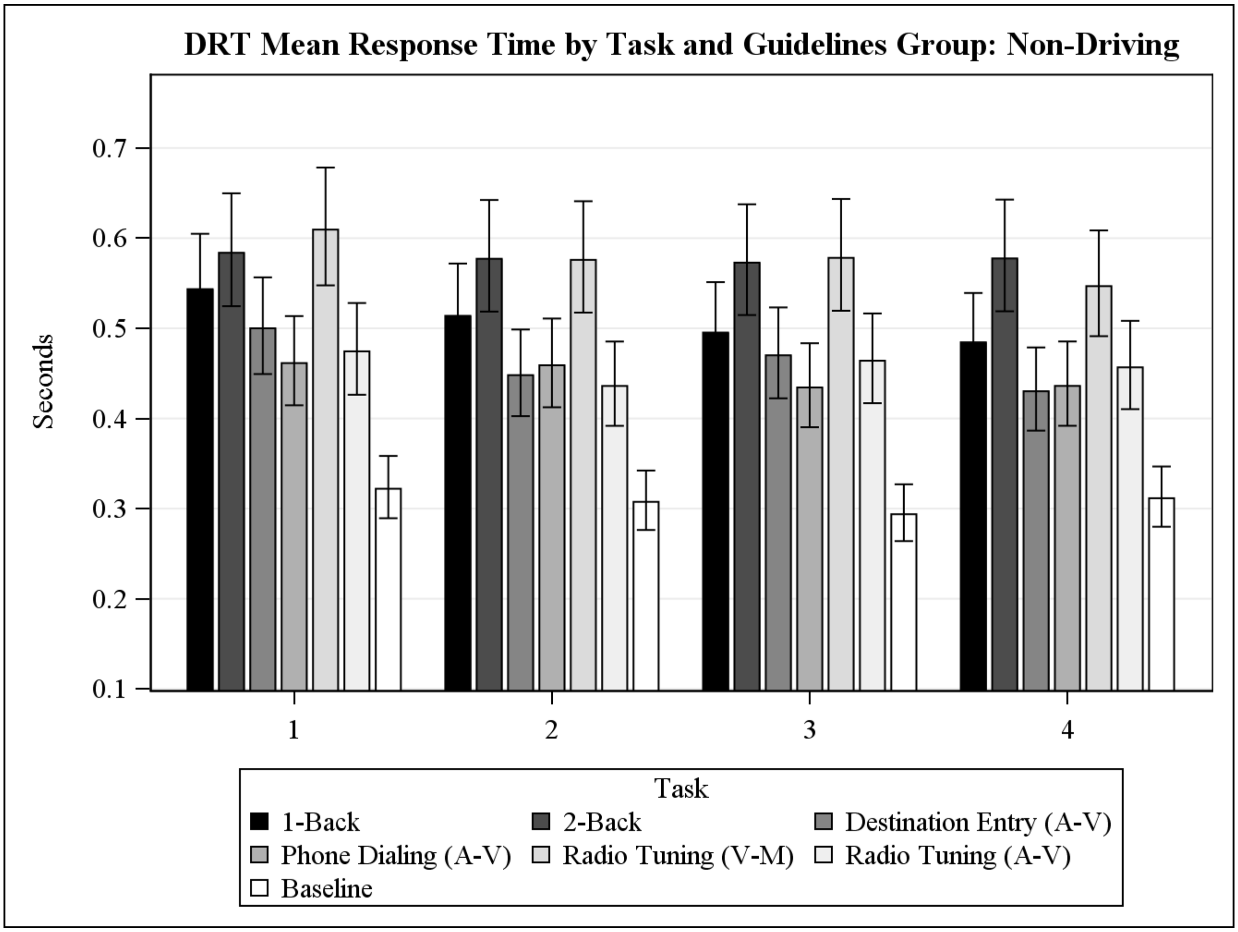


Figure 48. DRT Mean Response Time by Task and GL Group: Non-Driving

3.1.5.4 Non-Driving Detection Accuracy

Results of fixed-effect tests for the ANOVA with DRT accuracy for the non-driving venue are presented in Table 18.

Table 18. Summary of Fixed Effect Tests (Non-Driving Accuracy)

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**P > F**
Task	6	644	57.53	<.0001
Guidelines Group	3	116.8	0.34	0.7950
Guidelines Group*Task	18	644	1.91	0.0127

The GL Group x Task interaction effect is statistically significant. Means are presented below (Figure 49).

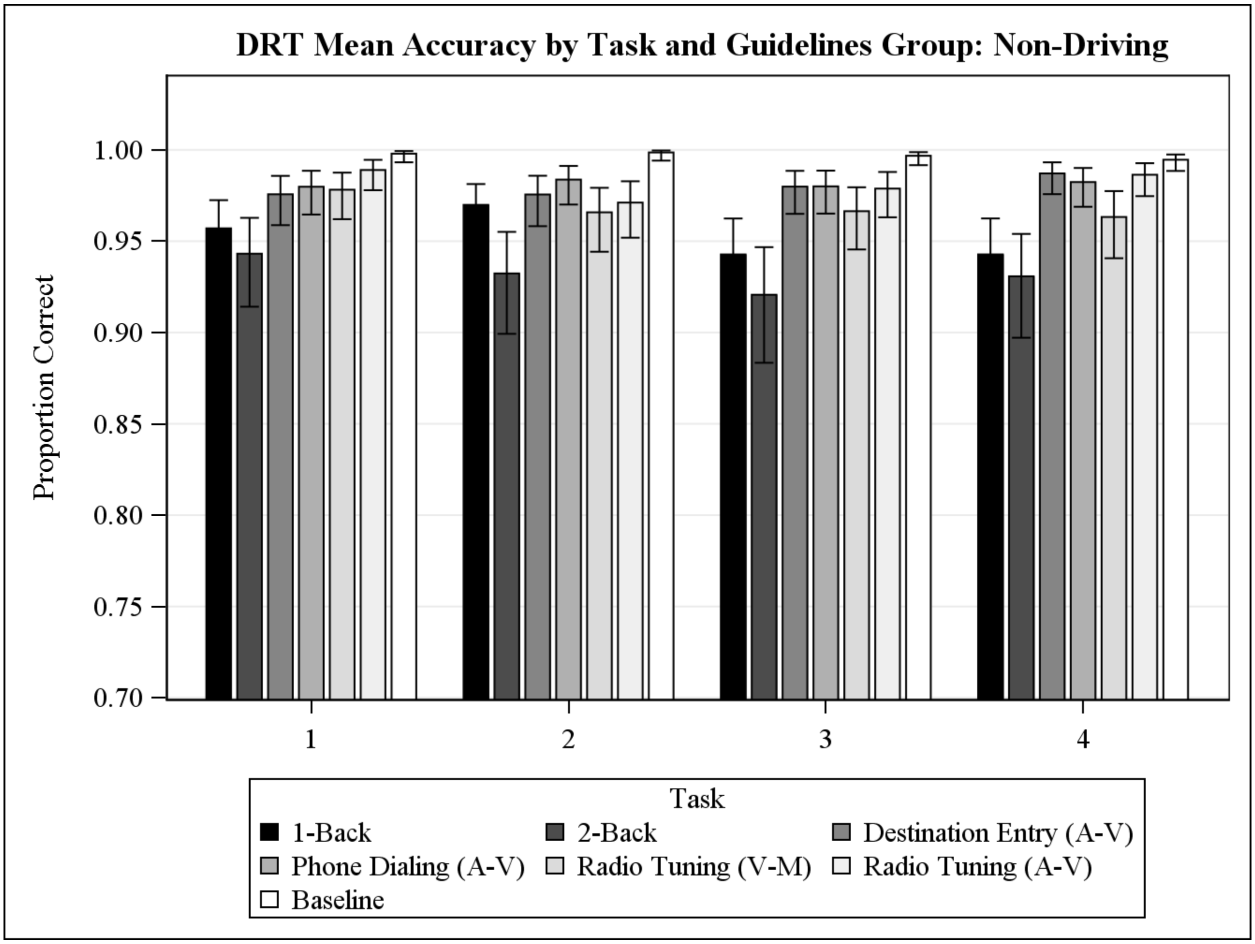


Figure 49. DRT Mean Accuracy by Task and GL Group: Non-Driving

The magnitude of differences between 1-back and 2-back differ among GL groups. The difference in accuracy between 1-back and the four in-vehicle tasks differs across groups. The 1-back mean is closer to the other task means for Groups 1 and 2 than for Groups 3 and 4.

3.1.6 Summary of DRT Findings

1. In the driving simulator, responses to TDRT targets were consistently slower, less accurate, and more variable than responses to RDRT targets. The slower responses together with reduced accuracy suggest that the tactile targets were more difficult to detect than the visual targets associated with the RDRT.
2. Differences in DRT RT and accuracy between DRT types, apparent in the driving simulator, are not present in the non-driving venue. DRT accuracy in the non-driving venue appears constrained by the ceiling of perfect performance, which may make it difficult to detect differences among task conditions.
3. Mean RT values were moderately negatively correlated with mean DRT accuracy values in both venues, indicating that participants were not trading speed for accuracy. Rather, faster responses were generally correlated with greater accuracy.
4. Differences between DRT conditions were apparent in the results of planned comparisons for both metrics in both venues. The differences were most apparent in the simulator for the response time metric, for which 6 of 9 comparisons had different outcomes for the two DRT conditions. These results reflected differences in the mean response times for the benchmark (1-back and 2-back) conditions in relation to the in-vehicle tasks, together with the overall slower and less accurate responses associated with the tactile condition in the simulator.
5. Differences between venues and DRT conditions were apparent. However, the design of the experiment had these as between-subjects' factors, which means that some of the observed differences may have been due to differences between samples. This may have implications for the expected consistency of test outcomes for testing that involves relatively small samples. The sample sizes used in this study were relatively large and are thus likely to understate the expected differences in variability associated with small-sample testing.
6. Age main effects were present in each analysis, reflecting the expected finding that response times increase and DRT accuracy decreases with increasing age, generally.
7. Age x Task interactions were found in both simulator metrics but not in the non-driving metrics. These effects suggest that ordering of metric values by task differ for the different age groups.
8. Differences among Guidelines Groups were apparent. Some were associated with differences between DRT types in the simulator. In the non-driving venue, differences between groups were less apparent.

3.1.6.1 Defining a Threshold Level of Acceptable Attentional Demand

One objective of this work was to determine how the results of this experiment could be used to establish a maximum level of acceptable attentional demand. Setting this threshold is considerably more difficult than setting a threshold for visual demand for several reasons. In the visual-manual realm, there were precedents based on the Alliance of Automobile Manufacturers' "Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems" (hereafter referred to as "AAM Guidelines," 2006) and there was more existing research and consensus about the use of radio tuning as a benchmark task. In the attentional/cognitive realm, there is less data and not nearly as strong a consensus about a maximum level of acceptable demand. Based on past related research

(Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014), it was concluded that the level of attentional demand represented by the 2-back task condition is unacceptable for tasks performed while driving. Similarly, 1-back was determined to represent a level of attentional demand that is acceptable for tasks performed while driving. The question to be explored in this section is whether these benchmark levels are sufficient for specifying a threshold of acceptable attentional demand that can be used generally to evaluate the attentional demand of various in-vehicle tasks.

If the benchmarks described above are accepted, they can be as anchors on a hypothetical scale of attentional demand. The logic of a possible test paradigm is outlined in the following points:

1. If a task has attentional demand that is greater than or equal to the demand associated with 2-back, that task shall be considered too demanding for performance while driving.
2. If a task has attentional demand that is less than or equal to the demand associated with 1-back, that task shall be considered acceptable for performance while driving.

The effectiveness of this paradigm is predicated on the sensitivity of the DRT metrics for differentiating between 1-back and 2-back levels of attentional demand. Table 19 summarizes the results of tests conducted in the present experiment to assess the differences between 1-back and 2-back for DRT RT and accuracy metrics.

Table 19. DRT RT and Accuracy by Venue and DRT Condition

Venue	**Response Time**		**Accuracy**	
	Remote	**Tactile**	**Remote**	**Tactile**
Simulator	0.0025	0.789 (NS)	0.0007	< 0.0001
Non-Driving	< 0.0001	0.0411	0.0804 (NS)	< 0.0001

For DRT RT, the RDRT was more sensitive than the TDRT to differences between the 1-back and 2-back conditions. For DRT accuracy, the TDRT was apparently more sensitive to differences between these two conditions than the RDRT. These differences may reflect differences among samples as different groups of participants were assigned to different venue/DRT conditions.

During the experiment, it was noticed that some participants appeared to have comparable difficulty performing both the 1-back and 2-back conditions. Some participants confused the two conditions, changing from 2-back to 1-back or vice versa in the middle of a trial. To examine the range of individual differences between these two conditions, two metrics were defined for each individual, one to represent the difference in response time and one to represent the difference in detection accuracy between the two benchmark conditions. For each participant, their individual mean response time in the 1-back condition was subtracted from their individual mean response time in the 2-back condition. Positive RT differences were obtained if RTs in the 2-back condition were slower than those in the 1-back condition. The same construction was used to compute the difference in accuracy, with the 1-back accuracy being subtracted from the 2-back accuracy. Because 1-back RT was expected to be faster than 2-back RT, the distribution of this metric was hypothesized to be mostly positive. Similarly, because 1-back accuracy was expected to be greater than 2-back accuracy, the distribution of this metric was hypothesized to be mostly negative. Figure 50 presents box plots of the distributions of RT differences by test venue and DRT condition.

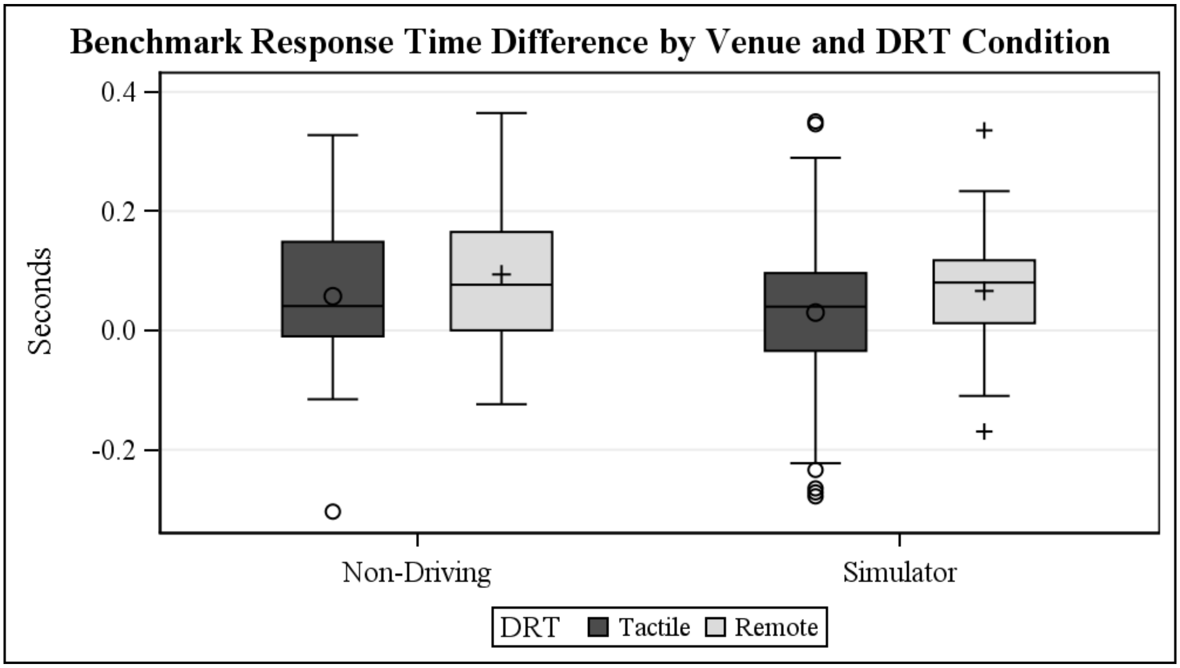


Figure 50. RT Difference (2-back vs. 1-back) Distributions by Test Venue and DRT Condition

As predicted, the differences between RT values were mainly positive, reflecting the typically slower RTs associated with the 2-back condition. However, at least 25 percent of each distribution was negative, reflecting longer mean RT values associated with 1-back than with 2-back. This pattern was most apparent for the TDRT in the simulator, which appears to have approximately 40 percent of differences contrary to the predicted direction. Recall that the difference between the 1-back and 2-back conditions was not statistically different for this condition. While most of each distribution reflected differences in the expected direction, the respective median values were relatively small: 0.08 (ND/RDRT); 0.04 (ND/TDRT); 0.08 (Sim/RDRT); and 0.04 (Sim/TDRT). Overall, the RDRT appears to be more sensitive to this difference and the distribution in the simulator had slightly fewer negative values.

Figure 51 presents the distributions of differences in DRT detection accuracy between the 1-back and 2-back conditions by venue and DRT condition.

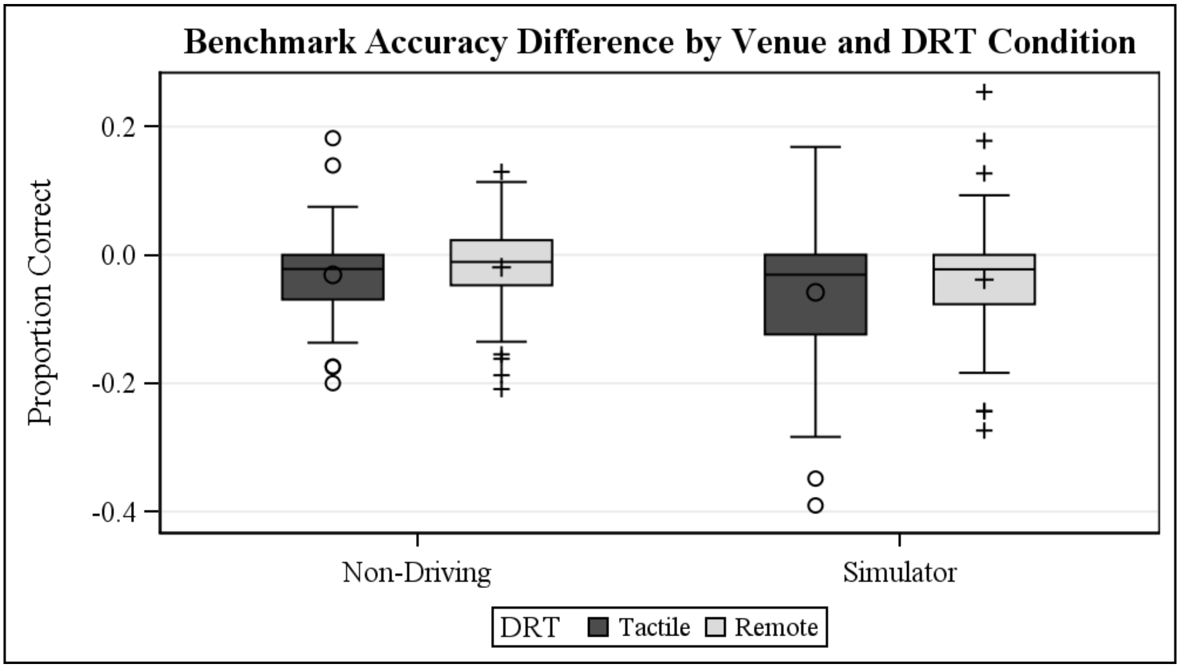


Figure 51. DRT Accuracy Difference (2-back vs. 1-back) Distributions by Test Venue and DRT Condition

For three of the four distributions, approximately 25 percent of differences were contrary to expectation, reflecting better DRT detection accuracy in the 2-back than in the 1-back condition. The fourth distribution (non-driving/remote DRT), had more than 25 percent of differences in the wrong direction. This observation is consistent with the result of statistical testing, which revealed the difference between 1-back and 2-back conditions to be not significantly different. As with the RT difference scores, the respective median values for these accuracy difference scores were relatively small: -0.01 (ND/RDRT); -0.02 (ND/TDRT); -0.02 (Sim/RDRT); and -0.03 (Sim/TDRT).

These presentations were undertaken to assess the feasibility of using individual DRT RT and/or accuracy values associated with 1-back and 2-back as benchmarks for assessing other tasks performed by the same individual. Use of individual mean RT values was intended to compensate for the relatively wide range of individual differences in reaction time among participants, particularly across age groups. For the RT metric, the TDRT in the simulator provided the largest proportion of participants with positive differences and the most consistency among participants. However, with 25 percent or more participants having no or minimal differences in DRT performance between the two benchmark conditions, the use of a combination of individual 1-back and 2-back condition scores does not seem feasible. However, it may still be feasible to consider defining test outcomes based on individual task DRT RT means in relation to the DRT mean RT in the 2-back condition. Based on RDRT results in the simulator, it appears that the overall 2-back mean RT is approximately 10 percent slower than the 1-back mean RT value. This percentage could serve as the basis for developing an individual criterion, according to which a person performing a task would be in conformance if his or her individual DRT mean RT value for the assessed task is at least 10 percent faster than his or her mean DRT RT from the 2-back condition. Use of group means and statistical testing is likely not feasible because a 10 percent difference may not be statistically significant for samples of N = 24.

3.1.6.2 Conclusions from DRT Analyses

The analysis results converge to indicate that testing in the non-driving venue does not always provide the same results as obtained in the driving simulator. It appears that the requirement for concurrent driving while performing in-vehicle tasks may have altered the way participants performed the tasks. When driving, performance is necessarily intermittent for tasks that require some interaction with the in-vehicle information system. DRT performance was slower and less accurate for these tasks in the simulator, reflecting the increased demands associated with intermittent performance. The apparent result of differential task performance in different venues is that the in-vehicle tasks that require interaction with the in-vehicle information system are more demanding in the driving simulator than in the non-driving venue relative to the benchmark tasks that have no interaction with an in-vehicle system.

In the driving context, participants appeared to give the tactile stimulus a lower priority than the visual stimulus. The fact that baseline detection speed and accuracy were comparable across DRT types indicates that the targets were equally salient when no task was performed.

The DRT mean RT value associated with the 2-back condition may be a candidate for assessing individual performance and classifying individuals' task performance into conformance or non-conformance groups. This would mean that the 2-back task would represent an unacceptable level of attentional demand for auditory-vocal tasks performed while driving and that tasks must be less demanding to be considered acceptable. Based on the overall differences between 1-back and 2-back RT values, acceptable tasks could possibly be defined as those with 10 percent better performance than 2-back. Adopting this value would require additional empirical support.

3.2 BRT Analyses

This section presents results pertaining to the sixth study objective.
Brake response time data was collected from 96 participants, each performing 4 driving trials with 7 braking events in each trial, totaling 28 braking events per participant and 2,688 braking events overall. In each trial, the first 6 braking events were expected responses to lead vehicle braking, indicated by brake light activation, while the seventh braking event was an unexpected situation in which the lead vehicle decelerated to a stop without using any brake lights.

During the unexpected emergency stop by the lead vehicle, the cues were the change in lead vehicle size and increased rate of closure while car following. Thus, the perception of these cues could be affected by how well the participant performed the car-following task at the time of event onset, which could in turn affect response time.

The brake channel data from all 96 participants was explored to determine an acceptable brake response criterion. In the simulator, the brake pedal input to the simulator was in the form of a range of input counts representing the displacement of the vehicle's brake pedal.

Calibration involved selecting minimum and maximum braking displacements and setting them to 0 and 100, respectively. Ten was selected, which represents 10 percent of maximum displacement as the criterion for defining the drivers' first brake response to the events. After examination of individual plots, it was decided that brake activity below the 10 percent threshold could not be reliably interpreted as a response to the lead-vehicle braking event, in which many values lower than 10 were affected by noise. Of the 2,688 braking events, this brake response time threshold provided 2,255 valid brake responses for analysis.

3.2.1 Brake Response Events: Expected and Unexpected Scenarios

Figure 52 shows the distribution of brake response times for both event types: the expected events in which the lead vehicle brake lights come on and the unexpected events in which the lead vehicle stops unexpectedly without any brake light activity. For the expected braking events (*N* = 1877), the mean brake response time was 1.290 s. For the unexpected braking events (*N* = 378), the mean brake response time was 1.925 s.

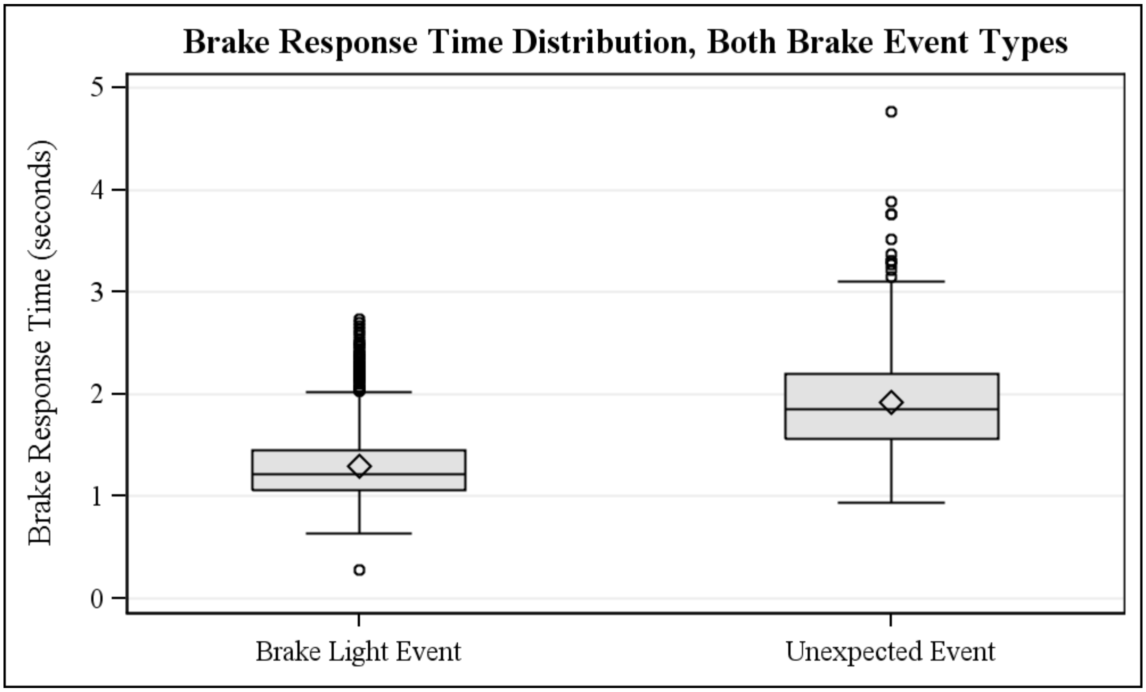


Figure 52. BRT Distribution, for Both Types of Braking Events

Table 20 provides a breakdown of the expected braking events (with brake lights on) by selected response time criteria, counting the number of response times occurring within each selected range of values.

Table 20. Summary of Number of Expected Braking Events by Ranges of Response Times

Expected Braking Events Summary	Counts
Less Than 2.5 s	1,869
2.5 to 3.5 s	8
3.5 to 4.5 s	0
4.5 s or more	0
No Valid Response	427
Total Count (96 x 6 x 4)	2,304

Expected Braking Events Summary	**Counts**
Maximum Response	2.732 s

After examination of individual plots, it was decided that brake activity below the 10 percent threshold could not be reliably interpreted as a response to the lead-vehicle braking event. For the 427 braking events listed in the table as not having a valid response, 416 of those events had some form of brake activity below the chosen threshold, in which the participant did not press the pedal enough to reach the 10 percent brake pedal application criteria used for determining brake response time. Only 8 of the 427 events with no valid response had no brake activity at all, and for 3 other braking events, the participants were already braking at or above the brake response criteria at the time when those braking events occurred. Over 99 percent valid responses occurred within the first 2.5 s following brake light onset.

Table 21 provides a breakdown of the unexpected braking events by selected response time criteria, counting the number of response times occurring within each selected range of values.

Table 21. Summary of Number of Unexpected Events by Ranges of Response Times

Unexpected Braking Events Summary	**Counts**
Less Than 2.5 s	332
2.5 to 3.5 s	41
3.5 to 4.5 s	4
4.5 s or more	1
No Valid Response	6
Total Count (96 x 1 x 4)	384
Maximum Response	4.765 s

For the 6 unexpected events listed in the table as not having a valid response, 5 of those events had some form of brake activity below the chosen threshold, meaning the participant did not press the pedal enough to reach the 10 percent brake pedal press criteria used for determining brake response time. Only 1 unexpected event had no brake activity at all, and there were no cases in which the participants were already braking at or above the brake response criteria at the time when those unexpected events occurred. Approximately 12 percent (46 of 378) of valid responses occurred later than 2.5 s following the beginning of lead-vehicle braking.

Figure 53 shows the mean brake response times in response to the lead vehicle brake light events for each of the four in-vehicle task conditions. For the 1-back task (*N* = 486), the mean brake response time was 1.23 s. For the 2-back task (*N* = 479), the mean brake response time was 1.26 s. For the baseline task (*N* = 480), the mean brake response time was 1.22 s. For the V-M radio tuning task (*N* = 432), the mean brake response time was 1.46 s. Longer brake response times in the V-M radio tuning condition likely reflect the fact that this task had V-M demands that required the driver to look away from the forward roadway view intermittently.

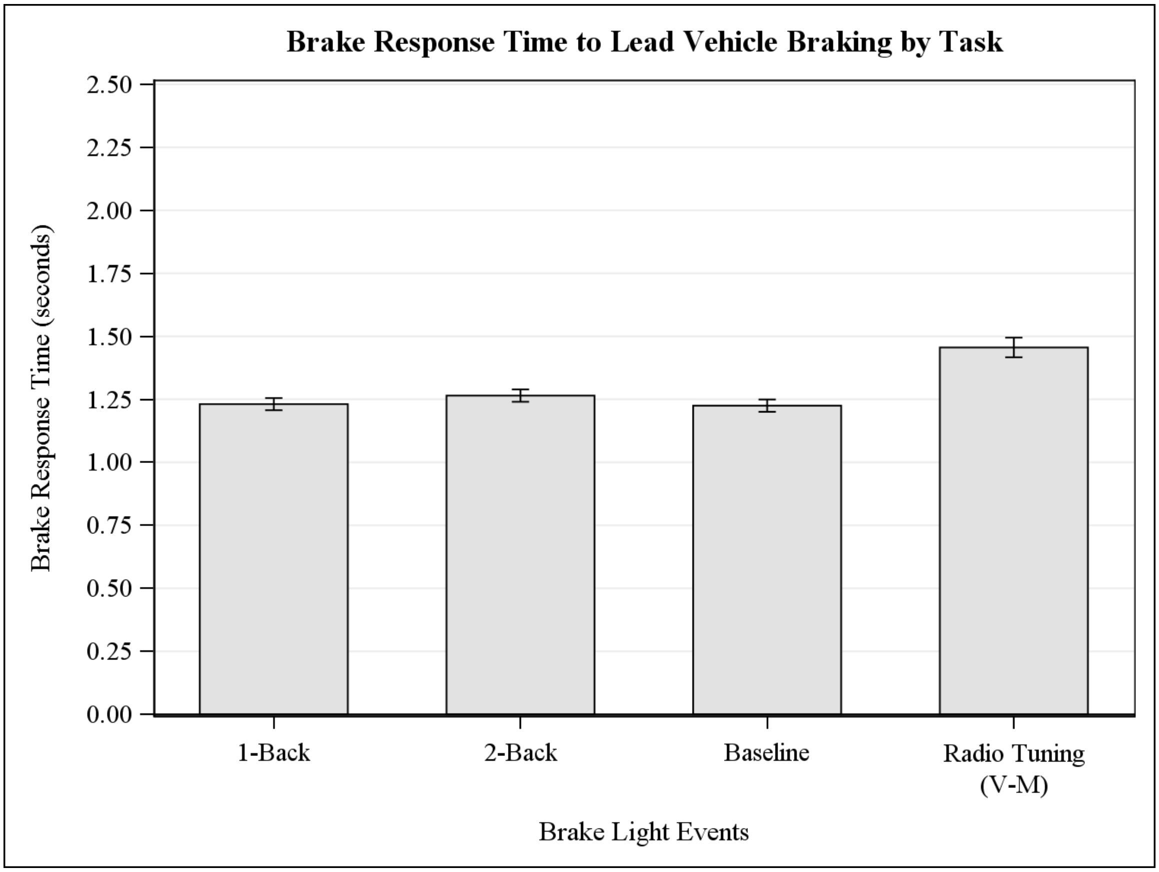


Figure 53. BRT to Lead Vehicle Brake Light Events by In-Vehicle Task

Figure 54 shows the mean brake response times of the V-M radio tuning trial, in which the lead vehicle brake light events are separated into those in which the V-M radio tuning task was active concurrently and those in which there was no active V-M radio tuning task at the time of brake light activity. There were 6 brake light events in each V-M radio tuning trial, in which: there were 3 (odd numbered) events where a V-M radio tuning task was active (Yes), and 3 (even numbered) events where the V-M radio tuning task was not active (No, just driving) during the trial. For when the V-M radio tuning task was active (N = 205), the mean brake response time was 1.544 s. For when the V-M radio tuning task was not active (N = 227), the mean brake response time was 1.376 s.

The other tasks were continuous effort during a drive, and did not have this active/not active status that was needed for V-M radio tuning.

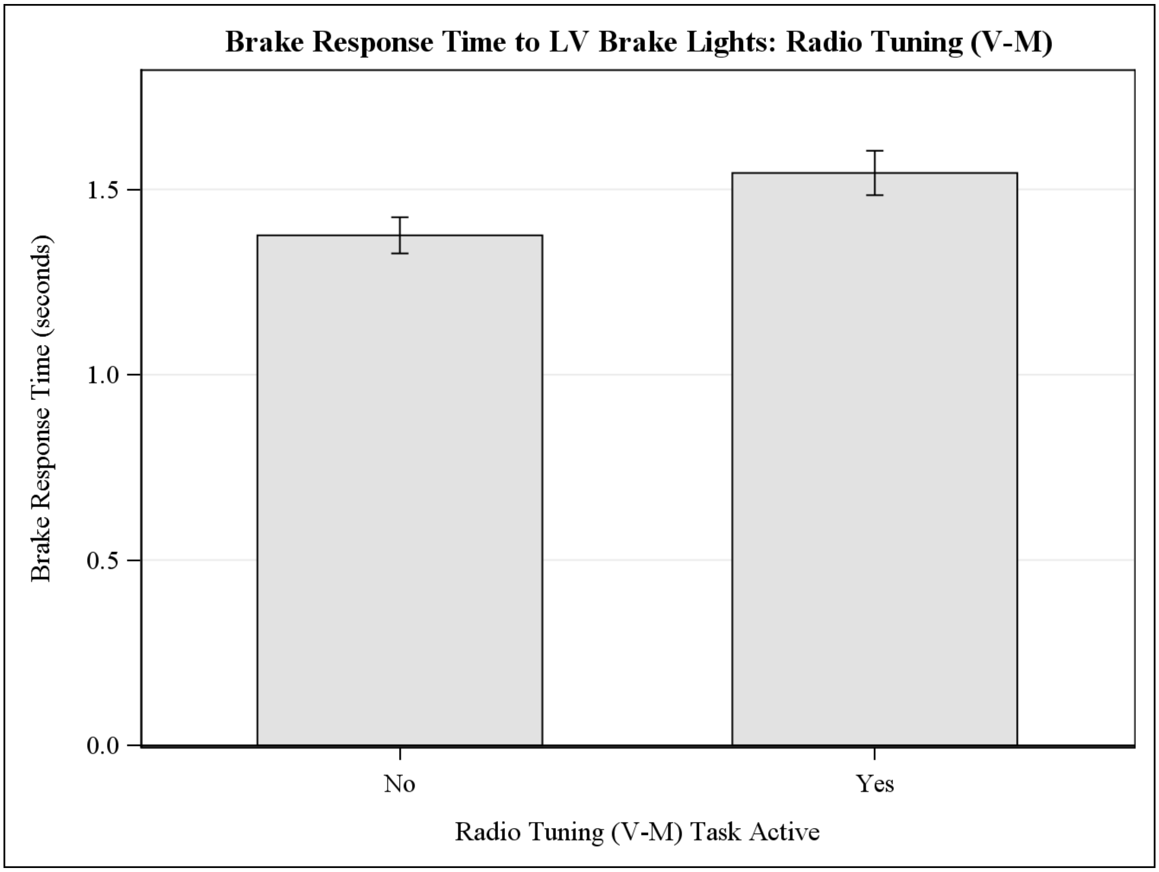


Figure 54. Mean BRT to LV Brake Lights for V-M Radio Tuning Trials

Figure 55 shows the mean brake response times in response to the unexpected lead vehicle stopping events for each of the four in-vehicle task conditions. For the 1-back task (*N* = 94), the mean brake response time was 1.859 s. For the 2-back task (*N* = 94), the mean brake response time was 1.901 s. For the baseline task (*N* = 95), the mean brake response time was 1.751 s. For the V-M radio tuning task (*N* = 95), the mean brake response time was 2.188 s.

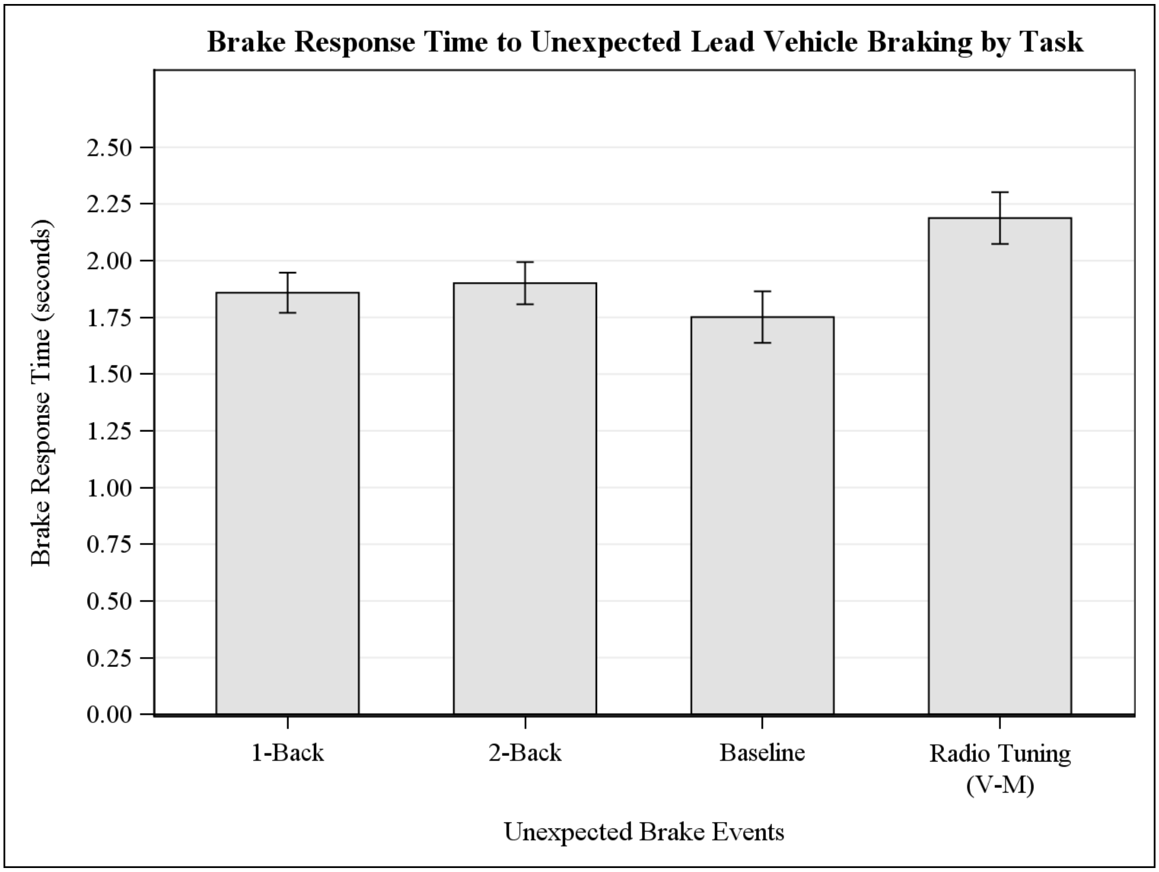


Figure 55. BRT to Unexpected Lead Vehicle Brake Events by In-Vehicle Task

Since the first unexpected event is more of a surprise event than the subsequent unexpected events, Figure 56 shows the details of just the first event by in-vehicle task, instead of all the unexpected events combined. For three of the four tasks, mean brake response time was higher during the first unexpected event than the mean response time for all the unexpected events combined. The V-M radio tuning task did not show this same trend. For the 1-back task (*N* = 24), the mean brake response time was 2.097 s. For the 2-back task (*N* = 23), the mean brake response time was 2.164 s. For the baseline task (*N* = 24), the mean brake response time was 1.926 s. For the V-M radio tuning task (*N* = 24), the mean brake response time was 2.144 s.

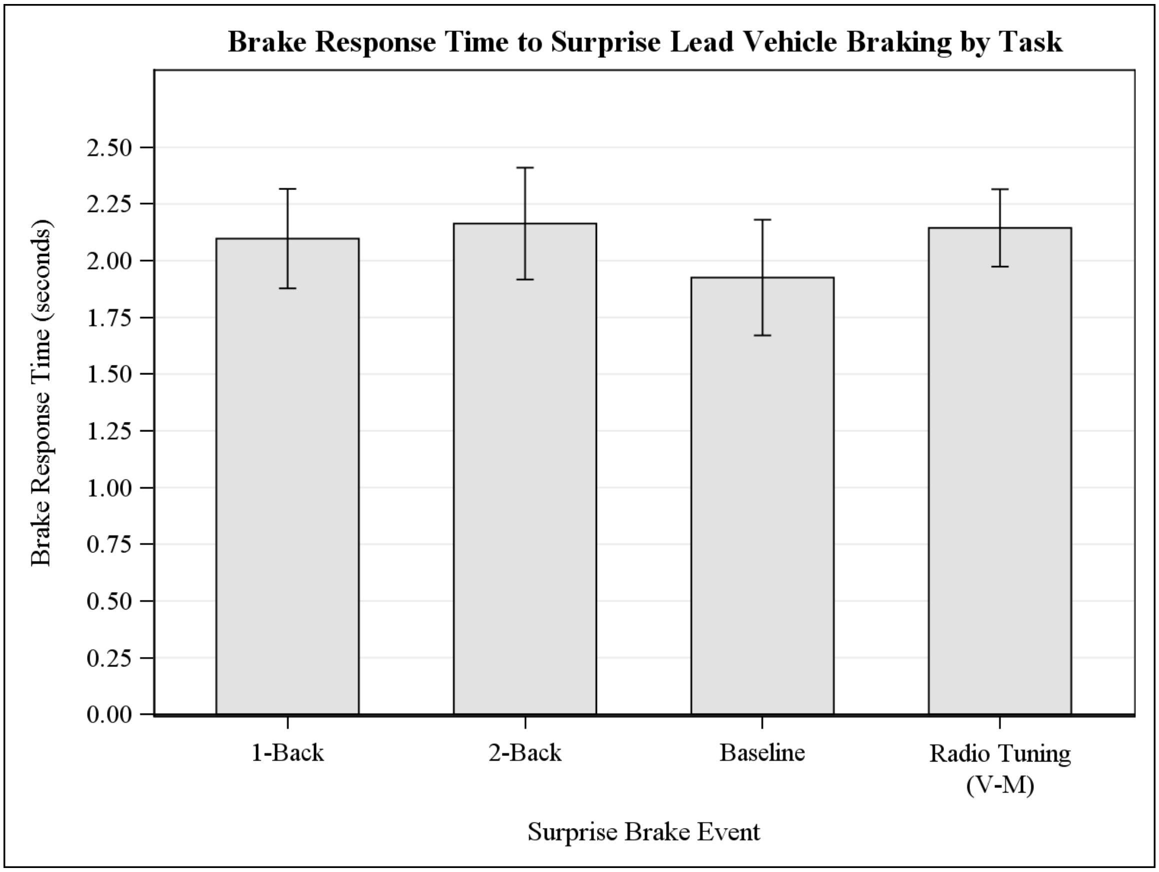


Figure 56. BRT to the Surprise (First Unexpected) Brake Event by In-Vehicle Task

The mean BRT for surprise events was faster in the baseline condition (*M* = 1.93 s, *SD* = 0.60 s) than in the 2-back condition (*M* = 2.16 s, *SD* = 0.57 s). Statistical testing was performed to determine whether the increase in attentional load between baseline and 2-back conditions was statistically significant. Parametric test results: *t*(45) = 1.45, *p* =.17, *d* = 0.40; Non-parametric: *t*(45) = 1.80, *p* = .08, *d* = 0.52.

In addition to brake response times to lead vehicle braking events, another metric of interest is headway, the simulated distance between the participant and the lead vehicle. Since whether a participant is correctly following at the specified car-following distance can influence responding to the lead vehicle brake events, the following Figure 57 presents the overall distribution of distance to the lead vehicle at the onset of the lead vehicle braking events, categorized by the two types of brake events experienced. For the lead-vehicle brake light condition (*N* = 2304), the mean distance to the lead vehicle was 113.7 feet at event onset. For the lead-vehicle unexpected stopping event condition (*N* = 384), the mean distance to the lead vehicle was 124.7 feet at event onset.

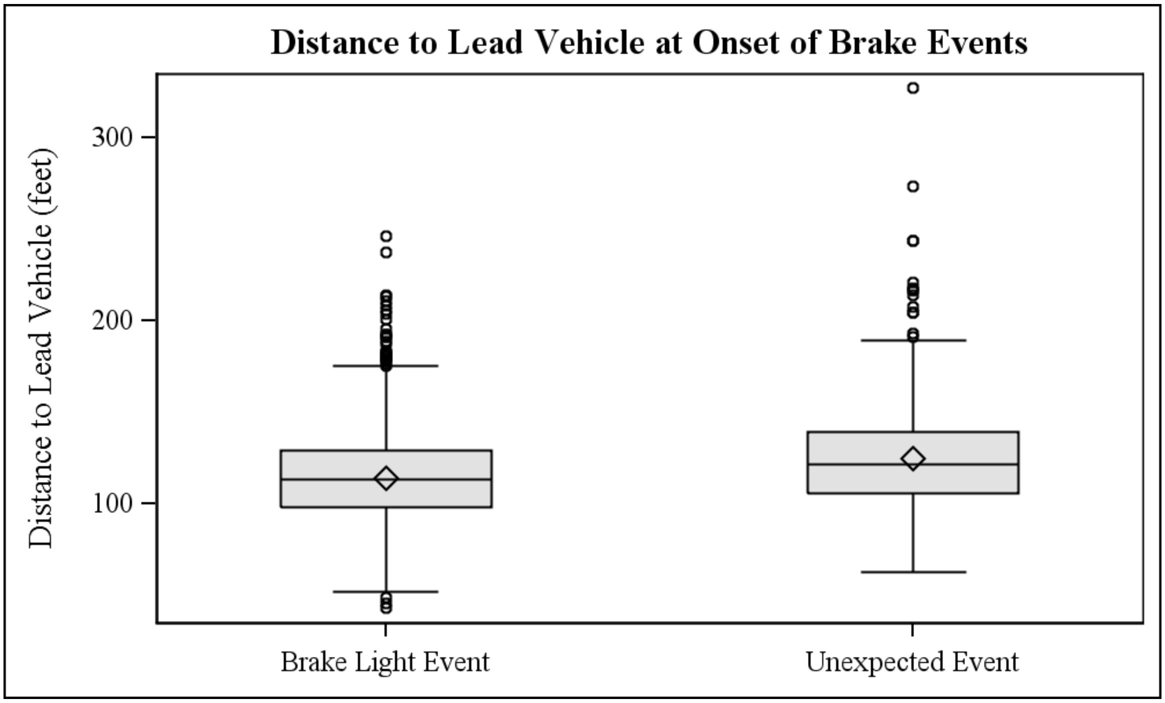


Figure 57. Distribution of Distance to Lead Vehicle at Brake Event Onset

3.2.2 Looming Cues in Emergency Scenarios

The "Distance to Lead Vehicle" is the simulator metric used to track car-following distance. Thus, when a participant is accurately performing the car-following task, this metric would match the specified 120-ft car-following distance at the onset of a lead vehicle's braking activity for both the expected brake-response events (with brake light activation) and the surprise events (no brake light activation). Since the surprise events did not have brake light activation, the participant would need to rely on looming cues (i.e., motion cues) to recognize an emergency response was required. In this scenario, the looming cues would be the rate of change in size (width) of the lead vehicle on the driver's retina, which depends on the size of the lead vehicle, the distance to the lead vehicle, and the closing speed between vehicles. It is estimated that drivers can detect changes in vehicle velocity when the angular velocity of the size of the vehicle on the retina is greater than approximately 0.003 radians/second (Hoffmann & Mortimer, 1996; Lamble, Laakso, & Summala, 1999).

To estimate angular velocity, the distance between the driver's eyes and the simulator screen was measured and the width of the lead vehicle on the screen at different distance-to-lead-vehicle positions computed by the simulator software. This data was plotted as shown in Figure 58. The data were fitted with the equation shown in the figure. This equation was used together with lead vehicle initial speed (73.33 ft/s), rate of deceleration (16.40 ft/s^2), and the distance of the driver from the screen (176 in., from an average driver's eye point location) to estimate the rate of change of the visual angle subtended by the vehicle in radians/second. These data are shown in the table below (Table 22).

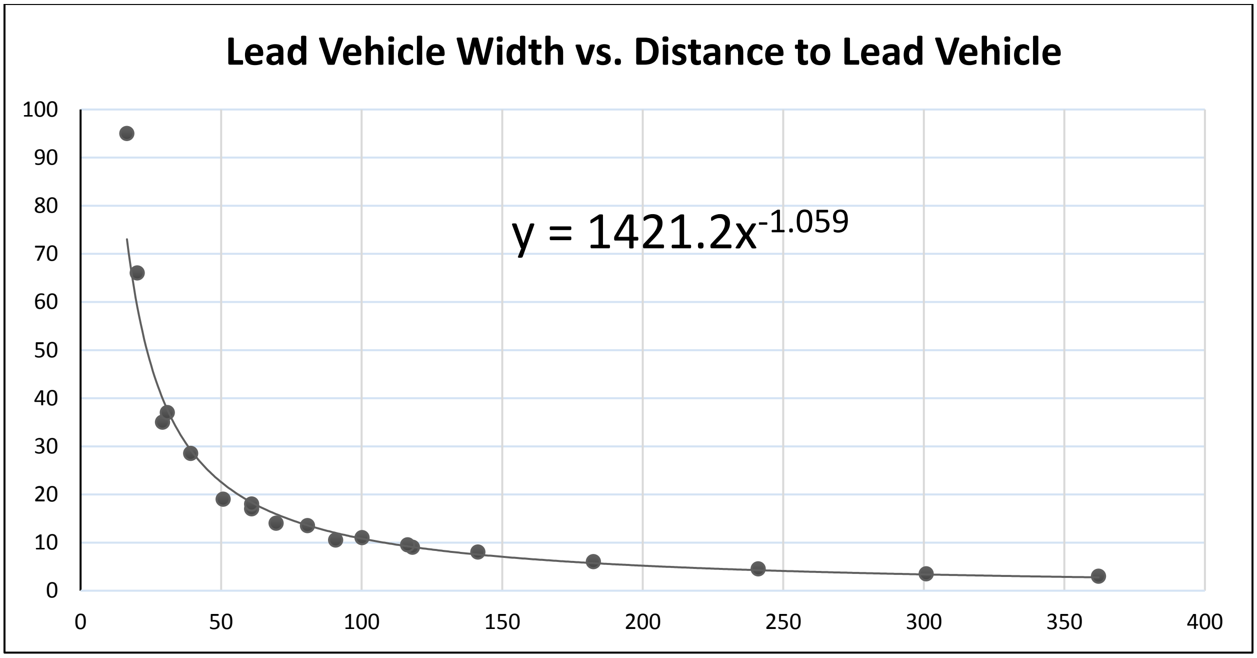


Figure 58. Relationship, On-Screen Vehicle Width and Simulated Distance to Vehicle

As shown in the final column in Table 22, the looming threshold (0.003 rad/s) is crossed when the following distance is approximately 116 ft.

Table 22. Example for Determining Rate of Change of Lead Vehicle

Time from Deceleration Event Start	(if Deceleration) Displacement (s), LV Decel	(if Constant Speed) Displacement, if 50 mph	New Car Following Distance (feet)	Vehicle Width (inch) $(y = 1421.2x^{-1.059})$	Distance to Screen (inch)	Visual Angle, Radians	Radian Diff	Time Diff	Radians / Second
0	**0**	**0**	**120**	**8.9**	**176**	**0.0507**	.	.	.
0.25	17.82	18.33	119.49	9.0	176	0.0510	0.0002	0.25	0.0009
0.5	34.62	36.67	117.95	9.1	176	0.0517	0.0007	0.25	0.0028
0.75	50.39	55.00	115.39	9.3	176	0.0529	0.0012	0.25	0.0049
1	65.13	73.33	111.80	9.6	176	0.0547	0.0018	0.25	0.0072
1.25	78.85	91.67	107.18	10.1	176	0.0572	0.0025	0.25	0.0100
1.5	91.55	110.00	101.55	10.7	176	0.0605	0.0034	0.25	0.0135
1.75	103.21	128.33	94.88	11.5	176	0.0650	0.0045	0.25	0.0180
2	113.86	146.67	87.19	12.5	176	0.0711	0.0061	0.25	0.0243
2.25	123.48	165.00	78.48	14.0	176	0.0795	0.0084	0.25	0.0335
2.5	132.07	183.33	68.74	16.1	176	0.0915	0.0120	0.25	0.0478
2.75	139.64	201.67	57.97	19.3	176	0.1095	0.0180	0.25	0.0722
3	146.18	220.00	46.18	24.5	176	0.1392	0.0297	0.25	0.1189
3.25	151.70	238.33	33.37	34.6	176	0.1961	0.0569	0.25	0.2276
3.5	156.19	256.67	19.52	61.1	176	0.3437	0.1475	0.25	0.5901
3.75	159.66	275.00	4.66	278.7	176	1.3393	0.9956	0.25	3.9824

The following figures explore the grouping of emergency brake response times into three groups based upon the simulated distance to the lead vehicle at emergency event onset. The three groups by car-following distance are: the Far group over 141 ft, the Near group under 99 ft, and the Normal group from 99 to 141 ft (normal, as in their car following was close to the specified 120 ft at emergency scenario onset).

Figure 59 shows the mean brake response times by secondary task and following distance at event onset for the first unexpected event, referred to as the surprise event. Group frequencies were: Far (*N* = 13); Near (*N* = 18) and Normal (*N* = 64). As can be seen in the figure, participants who were further away produced longer average response times than participants who were closer to the lead vehicle at surprise event onset, for each of the secondary task types. The extremely large confidence intervals associated with the 1-back and baseline conditions were due to the small frequencies associated with these conditions (*N* = 3).

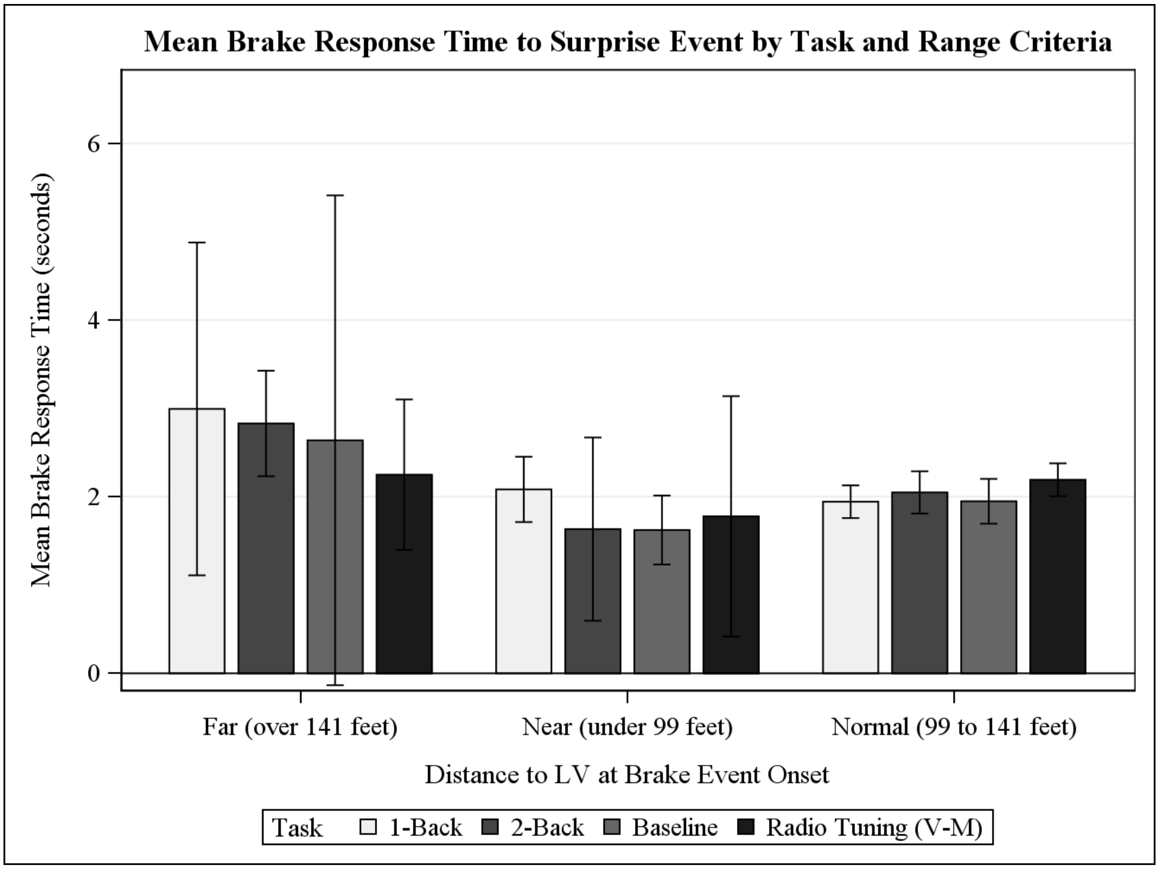


Figure 59. Surprise Event, Mean BRT by Secondary Task

Figure 60 shows the mean brake response times by secondary task and following distance at event onset for all unexpected events combined. Group frequencies were: Far (*N* = 84); Near (*N* = 64) and Normal (*N* = 230). As can be seen in the figure, the trend is the same as shown by the figure above of the first unexpected event, in which participants who were further away produced longer average response times than participants who were closer to the lead vehicle at event onset.

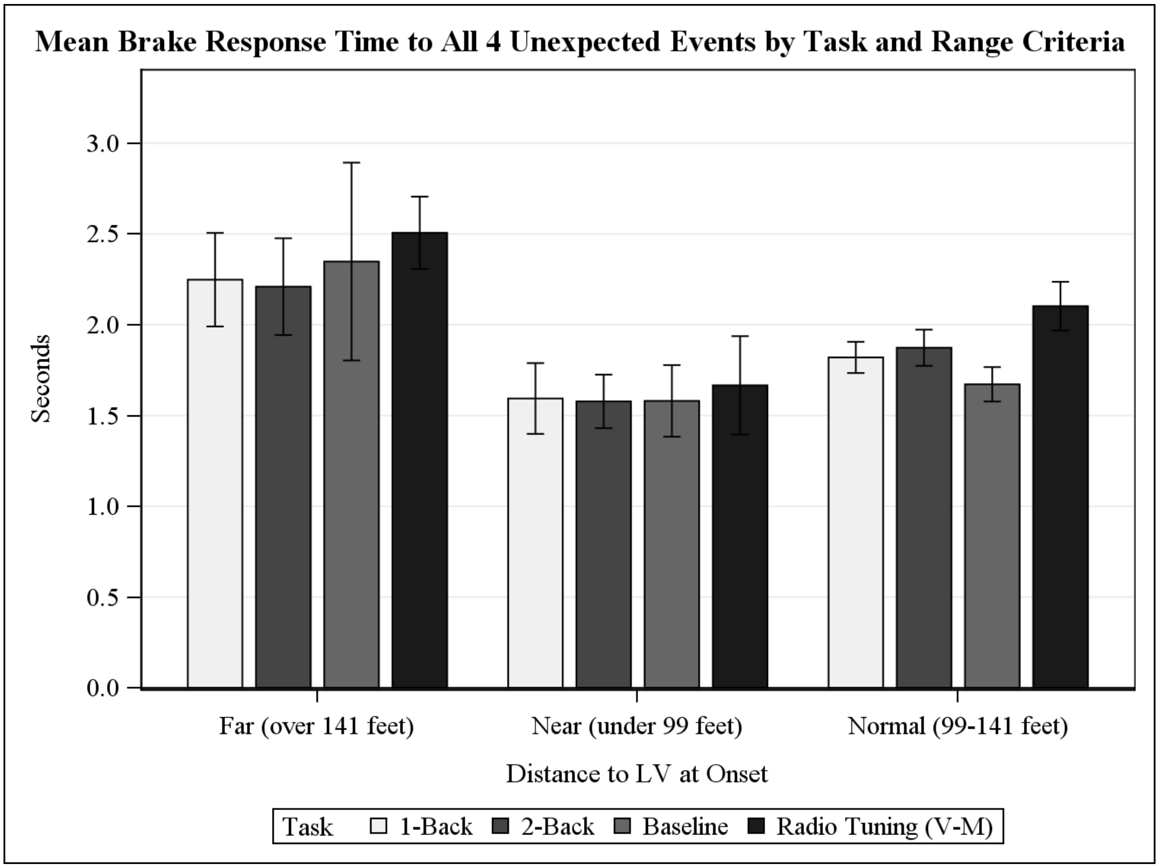


Figure 60. All Surprise/Unexpected Events, Mean BRT by Secondary Task

Figure 61 shows the mean brake response times by secondary task and following distance at event onset for the expected lead-vehicle brake light events. When compared to the figures above, following distance becomes less of a factor in response time when there are brake-light cues.

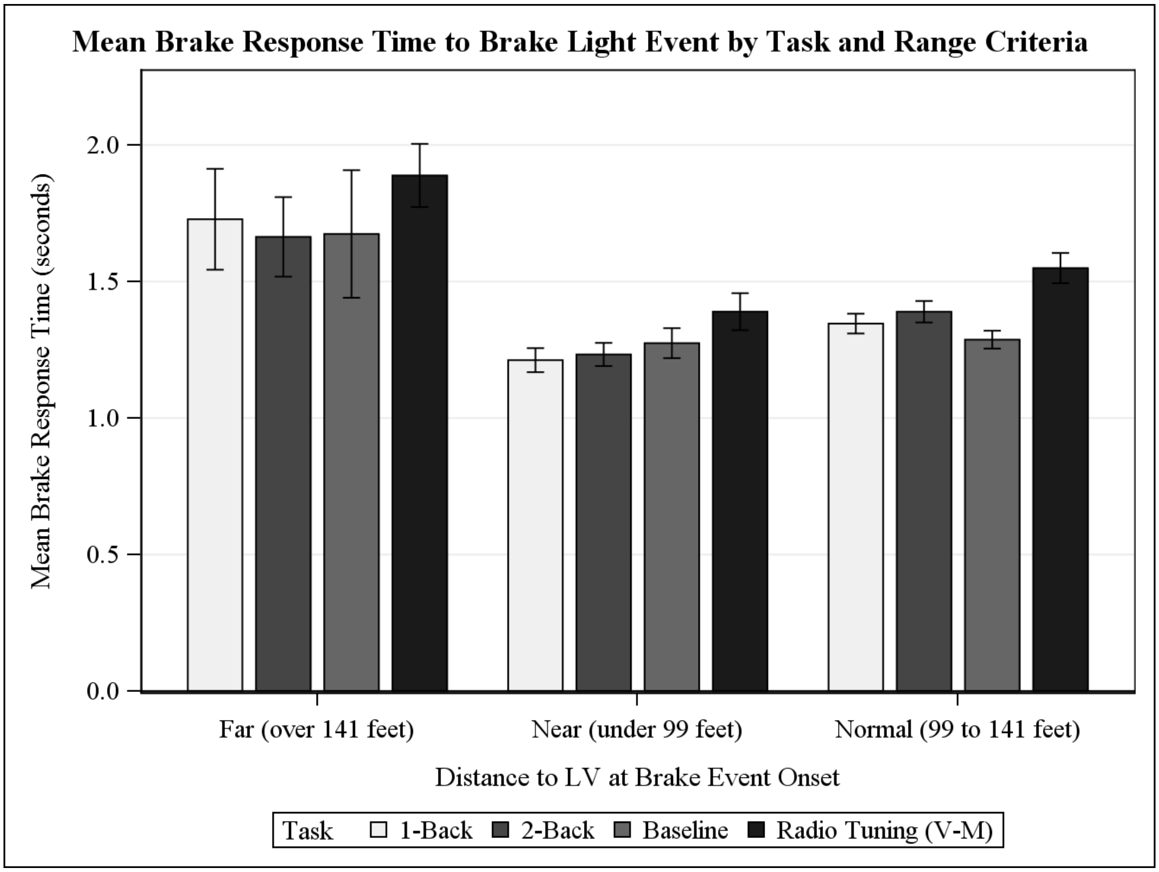


Figure 61. All LV Brake Light Response Events, Mean BRT by Secondary Task

Figure 62 is provided to show a more direct comparison of the mean brake response times for the two different brake response types, expected and unexpected, as well as show the trend when an unexpected event is used repeatedly.

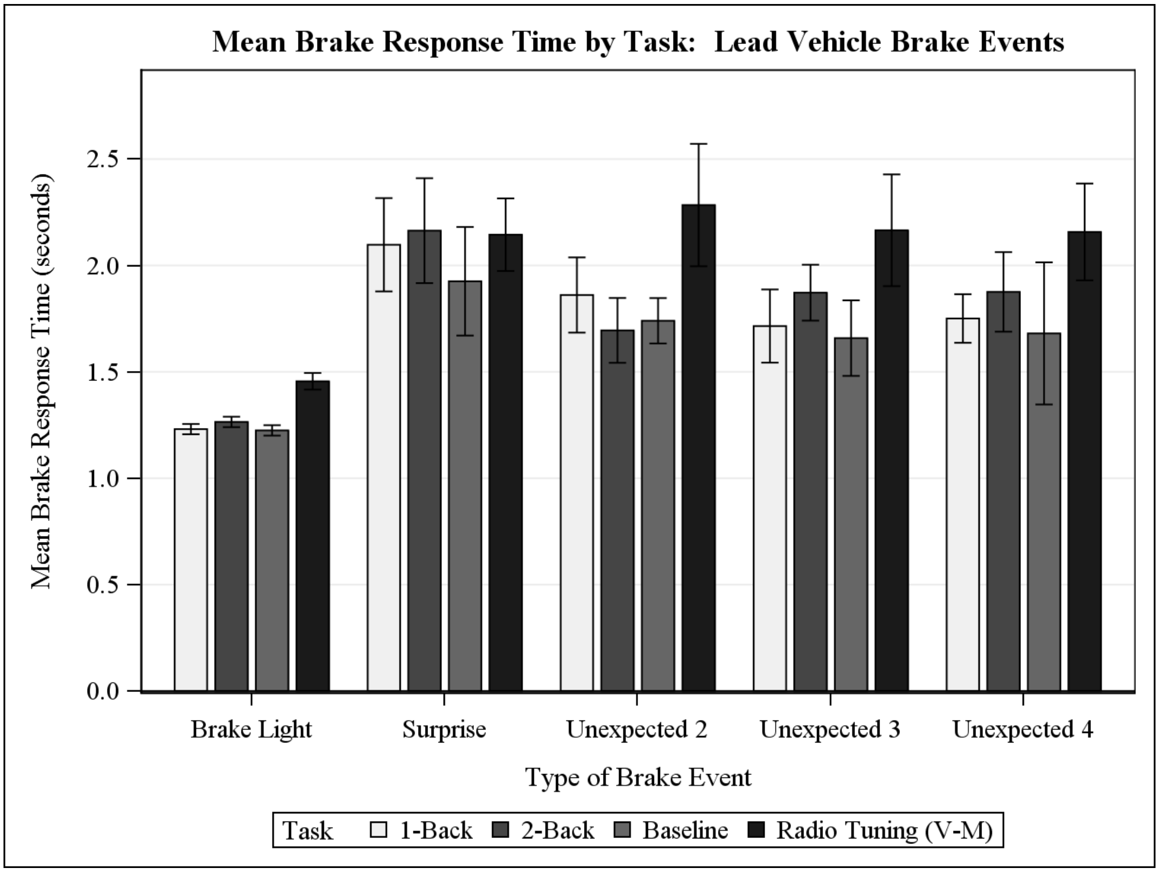


Figure 62. All Surprise and Brake Light Events, Mean BRT by Secondary Task

3.2.3 Summary of BRT Findings

1. As predicted, brake response times to expected lead-vehicle braking events were generally faster (*M* = 1.29 s) than responses to unexpected lead-vehicle braking events (no brake lights, *M* = 1.93 s).
2. When expected lead-vehicle braking events are considered in the aggregate, the overall mean BRTs associated with the 1-back, 2-back, and baseline conditions were approximately equal (1-back: 1.23 s; 2-back: 1.26 s; baseline: 1.22 s), suggesting that increased attentional load did not affect the brake response time to these expected events. The overall mean BRT associated with the V-M radio tuning condition was elevated relative to these conditions, reflecting a delay in responding associated most likely with the visual demands of the V-M radio tuning task.
3. In the V-M radio tuning condition, lead-vehicle braking tasks were scheduled so that half would occur when the participant was actively tuning the radio and half would occur when the participant was not actively tuning the radio. Although video data were not examined to establish the validity of the scheduling scheme, the mean response time during scheduled task performance was 1.54 s versus 1.38 s when performing the task was not scheduled. While the immediate attentional demand associated with the interval with no event scheduled was identical to baseline driving, the residual attentional demand associated with previous task performance is presumably responsible for the elevation of this mean above the 1.22-s baseline value.

4. Brake response time means associated with the unexpected events revealed small effects consistent with slowing due to attentional load. Specifically, the baseline mean BRT (1.75 s) was faster than the 2-back (1.90 s) and 1-back (1.86 s) means. The subset of these trials associated with the first unexpected event had slower responses, reflecting the increased uncertainty (surprise) associated with the first unexpected event (1-back 2.10 s; 2-back: 2.16 s; baseline: 1.93 s); however, the differences between conditions were only slightly greater.
5. Differences in brake response time between the baseline and 2-back conditions best reflect the effect of attentional load on brake response time. Table 23 summarizes the mean BRT values for these two conditions for the various event types in the experiment.

   Table 23. Mean BRT Values, Comparison of 2-back to Baseline Conditions

Condition	LV Brake (aggregate)	Unexpected (aggregate)	First Unexpected
2-back	1.26 s	1.90 s	2.16 s
Baseline	1.22 s	1.75 s	1.93 s
Difference (% relative to baseline)	0.04 s (3.2%)	0.15 s (8.6%)	0.23 s (11.9%)

   These results do not account for differences in the variability associated with the responses, however, the pattern of differences suggests that if there are any effects of attentional load, they exist in the group of brake responses to the first unexpected event, which had a 11.9 percent increase in BRT with the increased attentional demand. The much smaller difference between these two conditions observed in the aggregated LV brake data (0.04 s, 3.2%) suggests that responses to these expected events were not affected by attentional load. While the larger difference was not statistically significant, the effect size indicated a small effect.
6. The experiment did not allow for precise control of speed and headway at the start of the lead-vehicle braking events. To explore the effects of different headways at the start of the event, a set of analyses was conducted in which the events were separated by following distance at event onset. For the unexpected events, brake response times were influenced by headway; faster BRTs were associated with shorter headways. Among the subset of unexpected trials with headways close to the target value, BRT values appear to reveal an effect of attentional load with shorter BRTs associated with baseline and longer BRTs associated with 2-back. Among the subset of trials with shorter headways, there is no apparent effect of attentional demand, reflecting a pattern consistent with the predictions concerning looming effects, however, these trials are not all cleanly unexpected as discussed below. Visual cues associated with the decreasing distance to the lead vehicle had attained the looming threshold for both subsets.

3.2.4 BRT Discussion and Conclusions

Recent theories and research results have suggested that attentional load is likely to affect responses to events that are anticipated but not to those that are unexpected. This may be due to the engagement of cognitive resources in the anticipation of expected events, whereas no similar attentional activity would be involved before events that are not expected. For unexpected events, drivers’ responses are thought to be triggered primarily by looming cues associated with closing distance and increasing apparent size of the lead vehicle. Specifically, researchers have estimated that the threshold for detection of a closing object is approximately 0.003 radians/second (Lamble, Laakso, & Summala, 1999; Hoffmann & Mortimer, 1996). Accordingly, it is argued that human response to looming cues is a hard-wired survival mechanism that is not affected by attentional load. These theoretical positions lead to the following specific predictions in this experiment.

1. Brake response times to expected lead-vehicle braking events are predicted to increase with increasing attentional load.
2. Brake response times to unexpected lead-vehicle braking events, discernible only from looming cues, are predicted to be unaffected by attentional load.

When taken together, the present results were not consistent with either of these predictions. BRTs in the expected trials showed no effect of cognitive load while those associated with the first unexpected event showed a modest increase with increasing attentional demand. However, when the unexpected brake event trials were separated by headway at time of event onset, different patterns of responses emerged by task condition. For the first unexpected event, which represents the set of truly unexpected events, none of the distance groups revealed BRT patterns that were consistent with either the attentional load model (i.e., 2-back slower than baseline) or the looming model (i.e., no difference among 1-back, 2-back, and baseline). However, for the full collection of unexpected events, the subset of trials with shortest headways did reveal a pattern of means consistent with the looming model. Specifically, BRT values were approximately equal for 1-back, 2-back and baseline conditions. The headway values of all trials in this group were above the looming threshold defined above, however, it is reasonable to question whether the later unexpected events were truly unexpected and thus likely to be unaffected by drivers' anticipation. The pattern of differences associated with the middle-distance group was different; BRT in the baseline condition was faster than for 1-back or 2-back, suggesting an effect of attentional load for drivers at this distance. For the drivers in this group who had longer following distances at the time of the event, there is some uncertainty about whether looming cues were effective immediately. Thus, while one subset of data provided a pattern that could be interpreted to be consistent with the second hypothesis above, the between-subjects' design and the inclusion of events that were possibly not entirely unanticipated precludes a strong conclusion. Most generally, the pattern of BRTs was not supportive of either of the hypotheses presented above.

3.3 Occlusion Analyses

This section presents results pertaining to the third and fifth study objectives.

The 96 participants assigned to the non-driving venue performed four in-vehicle tasks using the occlusion protocol specified in the NHTSA Distraction Guidelines. This protocol used alternating 1.5-s intervals in which the occlusion goggles were occluded (i.e., closed) and unoccluded (i.e., open). Participants completed five trials for each task. The performance metric is the total amount of unoccluded time, also called total shutter open time (TSOT) during task performance. Ultimately, TSOT would be compared to total eyes off road time (TEORT) from the eye tracker in the driving simulator venue.

TSOT can be computed in two ways: a direct measurement and an open-interval measurement. The direct-computation method uses the total amount of shutter-open time between the beginning of the first open interval and the experimenter's recording of task completion.[5] The open-interval method for computing TSOT uses the total time associated with the number of open (i.e., unoccluded) intervals needed to complete the task. TSOT values computed using the open-interval method were on average 0.538 s greater, and up to 0.767 s greater than the corresponding TSOT values based on the direct-computation method. The open interval computation method was used because it was easier to compute and more consistent across trials.[6]

---

[5] The total duration of the interval (open plus closed intervals) is divided by two to obtain the TSOT.

[6] The longer TSOT values associated with the open interval method derive from the inclusion of full open intervals with part of the interval occurring after the task was completed. This could explain why the open interval method is consistently longer than the direct computation method.

3.3.1 Occlusion Analyses by Task

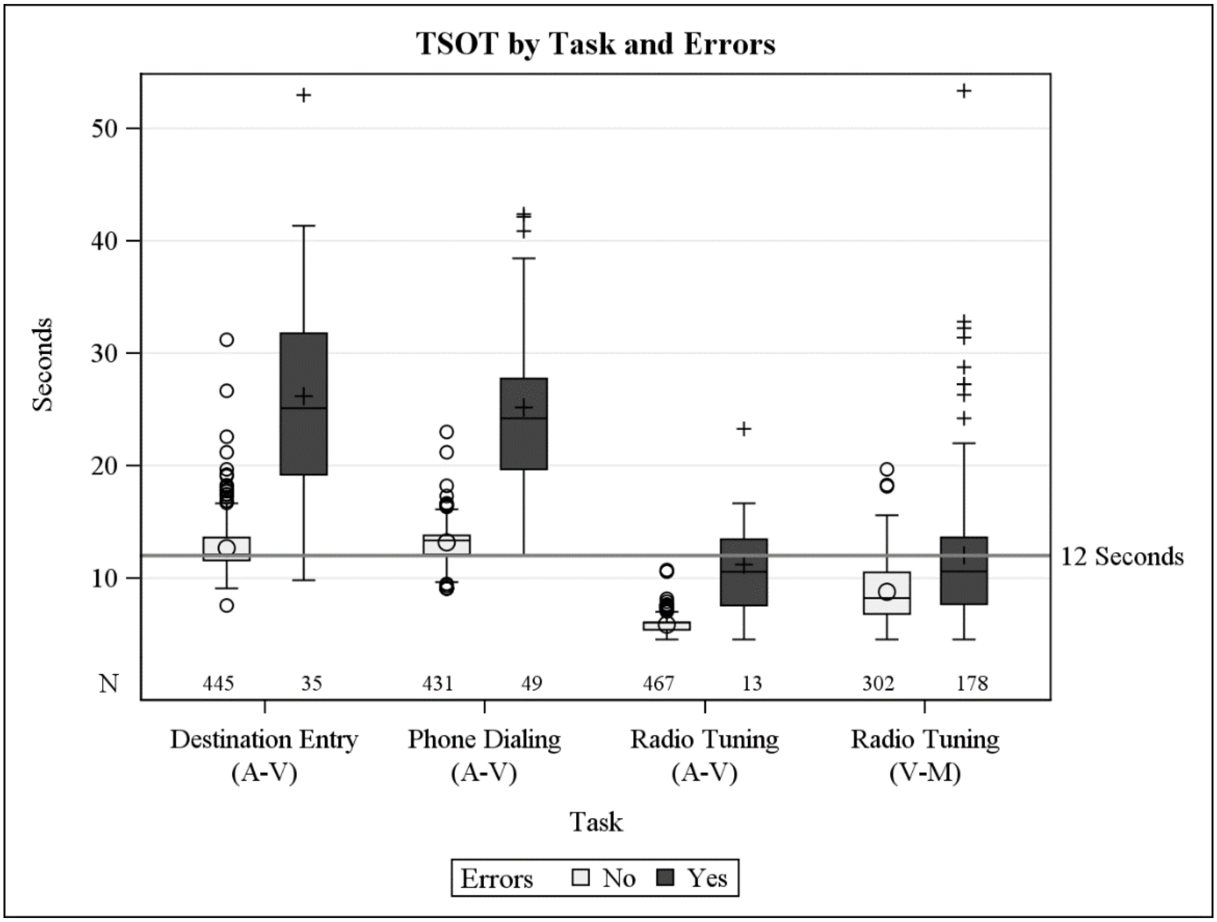


Figure 63 presents the boxplot distributions of TSOT for correct and error trials separately for the four tasks. Frequencies associated with each distribution are presented above the x-axis.

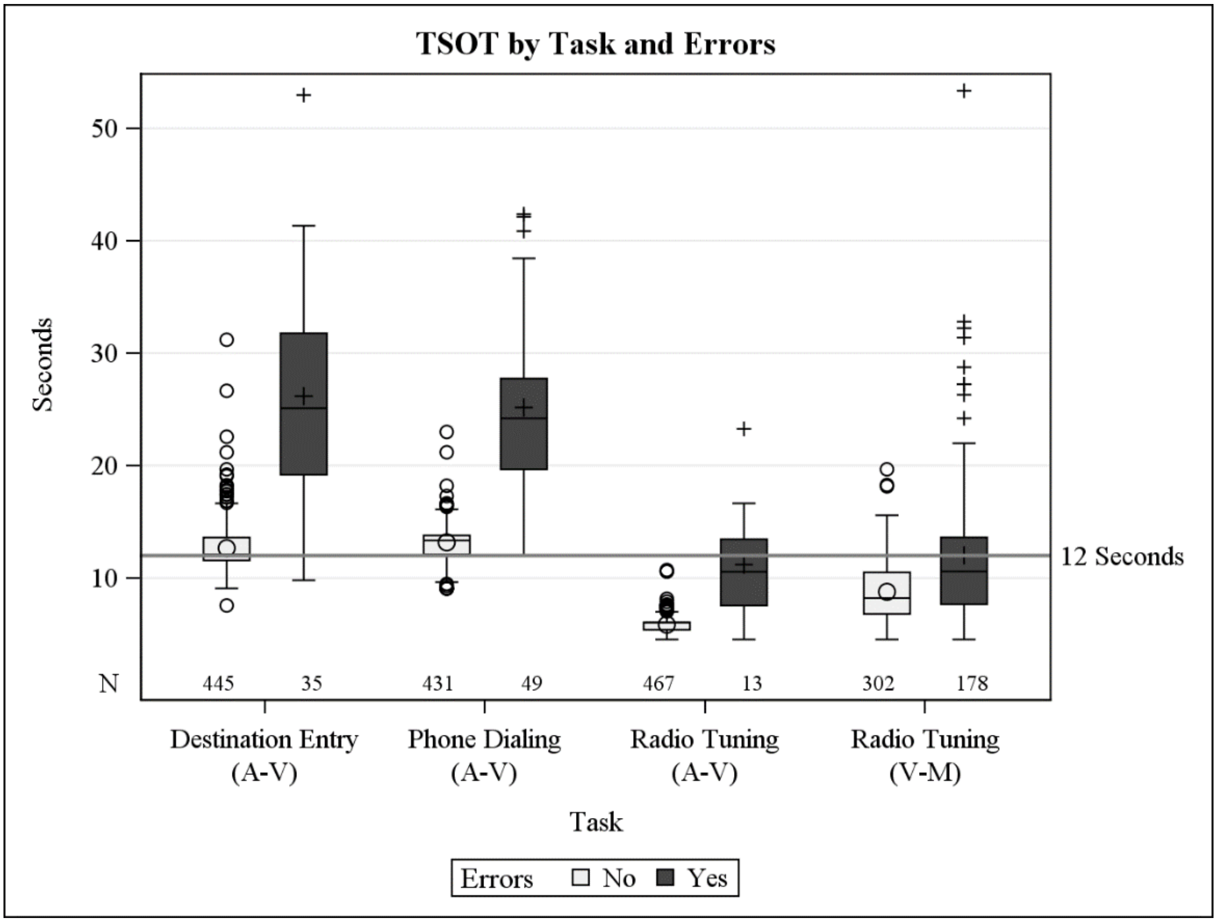


Figure 63. Occlusion TSOT Distributions for Correct and Error Trials by Task

Error proportions were: 0.073 for destination entry; 0.371 for V-M radio tuning; 0.102 for phone dialing; and 0.027 for A-V radio tuning. The distributions reveal that participants were generally unable to complete the destination entry and phone dialing tasks in the allotted 12 s of TSOT when an error was made. However, errors were not as disruptive, in terms of TSOT, for the two radio tuning tasks.

Additionally, the most errors occurred during the V-M radio tuning task. Figure 64 shows the distributions of mean TSOT, with one value computed for each task for each participant. For these distributions, participants with three or more error trials during task performance were not used to compute the mean values. Frequencies associated with each distribution are presented above the x-axis. Thirty-seven (0.385) V-M radio tuning participants and 5 (0.052) phone dialing participants were removed because they had three or more error trials.

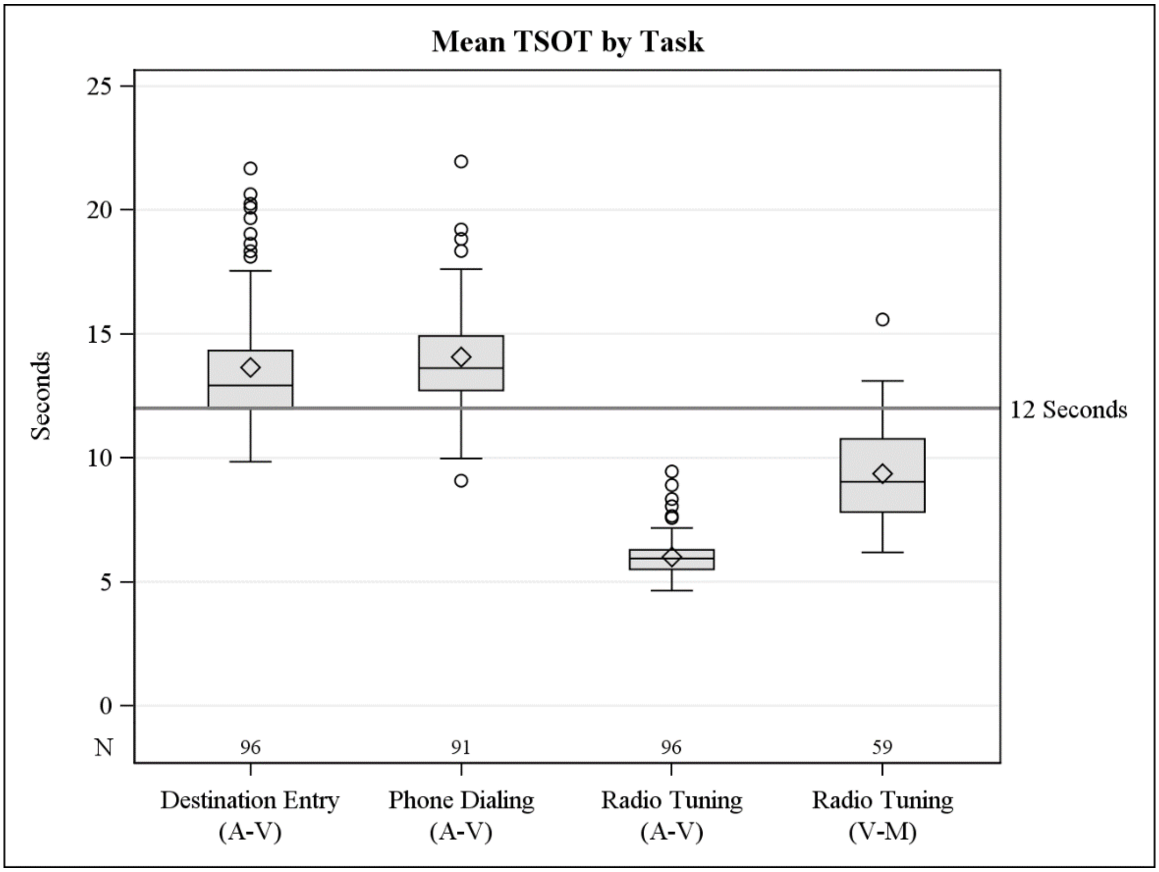


Figure 64. Mean TSOT Distributions by Task

Based on the aggregate data, it seems that the destination entry (*M* = 13.65, *SD* = 2.57) and Phone Dialing (*M* = 14.07, *SD* = 2.09) tasks are unlikely to be in conformance with the 12-s TSOT reference time. Both tasks were A-V tasks and appear to have required too many glances to complete the task in less than 12 s of TSOT. The two radio tuning tasks appear to be in conformance with this criterion (V-M radio tuning: *M* = 9.36, *SD* = 1.91; A-V radio tuning: *M* = 6.0, *SD* = 0.83).

Figure 65 presents the distributions of mean TSOT separated by age group and task condition. Age effects, reflected in increasing TSOT values with increasing age, are apparent for the V-M radio tuning task and to a lesser extent for the phone dialing task. Among the younger participants (i.e., ages 18-24), destination entry was associated with considerably more variability among mean TSOT values than for the other tasks.

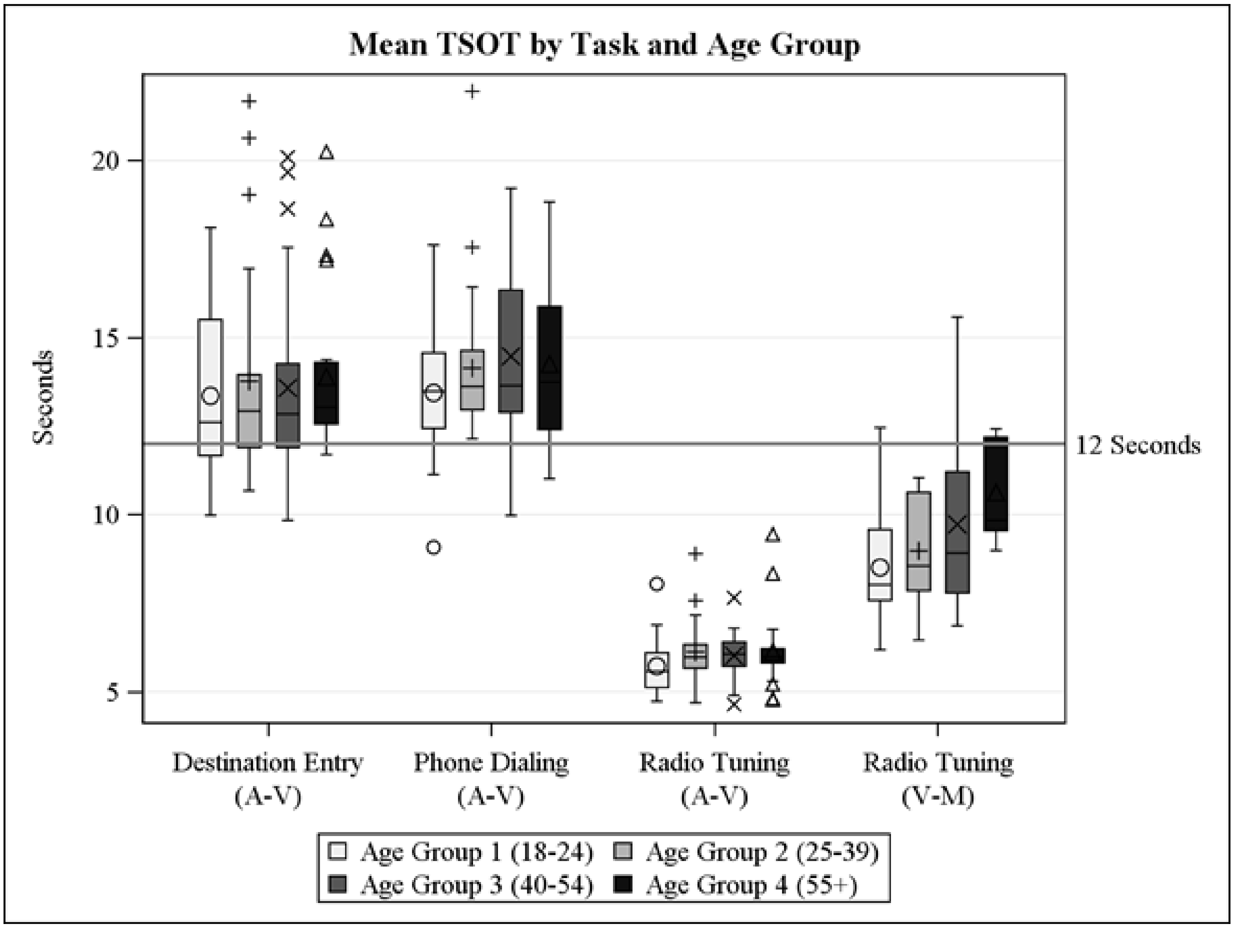


Figure 65. Mean TSOT Distributions by Task and Age Group

3.3.2 Occlusion Analyses Based on Guidelines Test Criterion

The NHTSA Driver Distraction Guidelines specified 12 s as the maximum TSOT allowed while performing a task. The 96 participants who performed occlusion trials comprised four 24-person Guidelines samples based on the age and gender requirements presented therein. Distribution box plots for the mean TSOT for each of the Guidelines samples and tasks are presented in Figure 66. With respect to the criterion value, the groups were relatively consistent. Groups varied in the amount of variability; however, these differences between groups were not consistent across task conditions.

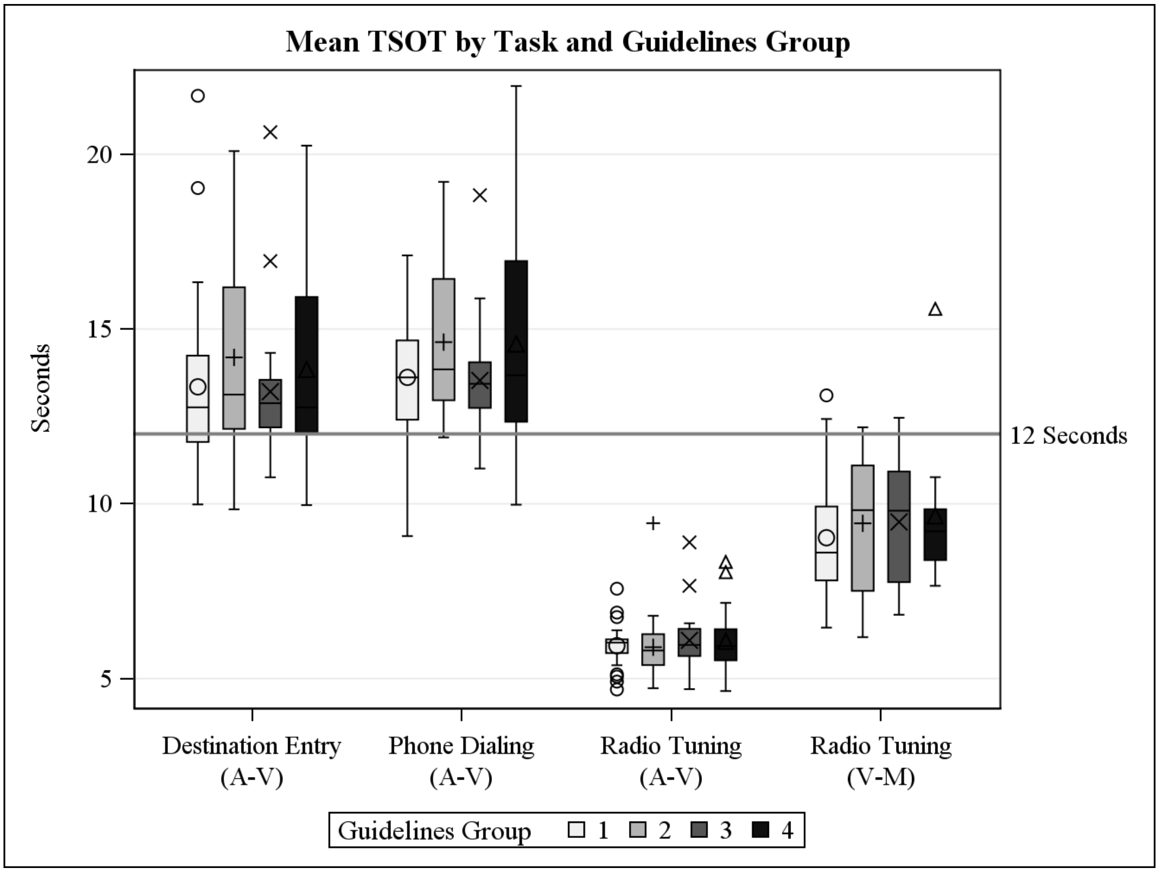


Figure 66. Mean TSOT Distributions by Guidelines Sample and Task

3.3.3 Conformance to Guidelines TSOT Criterion

3.3.3.1 Conformance by Task

Looking further at the TSOT based on NHTSA Driver Distraction Guidelines test criterion, Figure 67 shows the distributions of individual mean TSOT values above and below the Guidelines 12-s TSOT allowed for each task. Each bar represents the mean TSOT value for one participant per task after removing participants with more than two error trials. For the following figures, participants who did or did not complete the task within the TSOT criterion were indicated as conforming or non-conforming, respectively. The 12-s TSOT criterion value is indicated with a reference line.

The two radio tuning tasks had the most participants who were in conformance, with 100 percent of A-V radio tuning participants (*N* = 96) and 88.1 percent of V-M radio tuning participants (*N* = 59) in conformance. Destination entry and phone dialing had the least participants that were in conformance with the 12 s TSOT reference time: 27.1 percent of destination entry participants (*N* = 96) and 7.7 percent of phone dialing participants (*N* = 91) were in conformance. These findings are also shown in the boxplots of the TSOT means for each task (above, Figure 64).

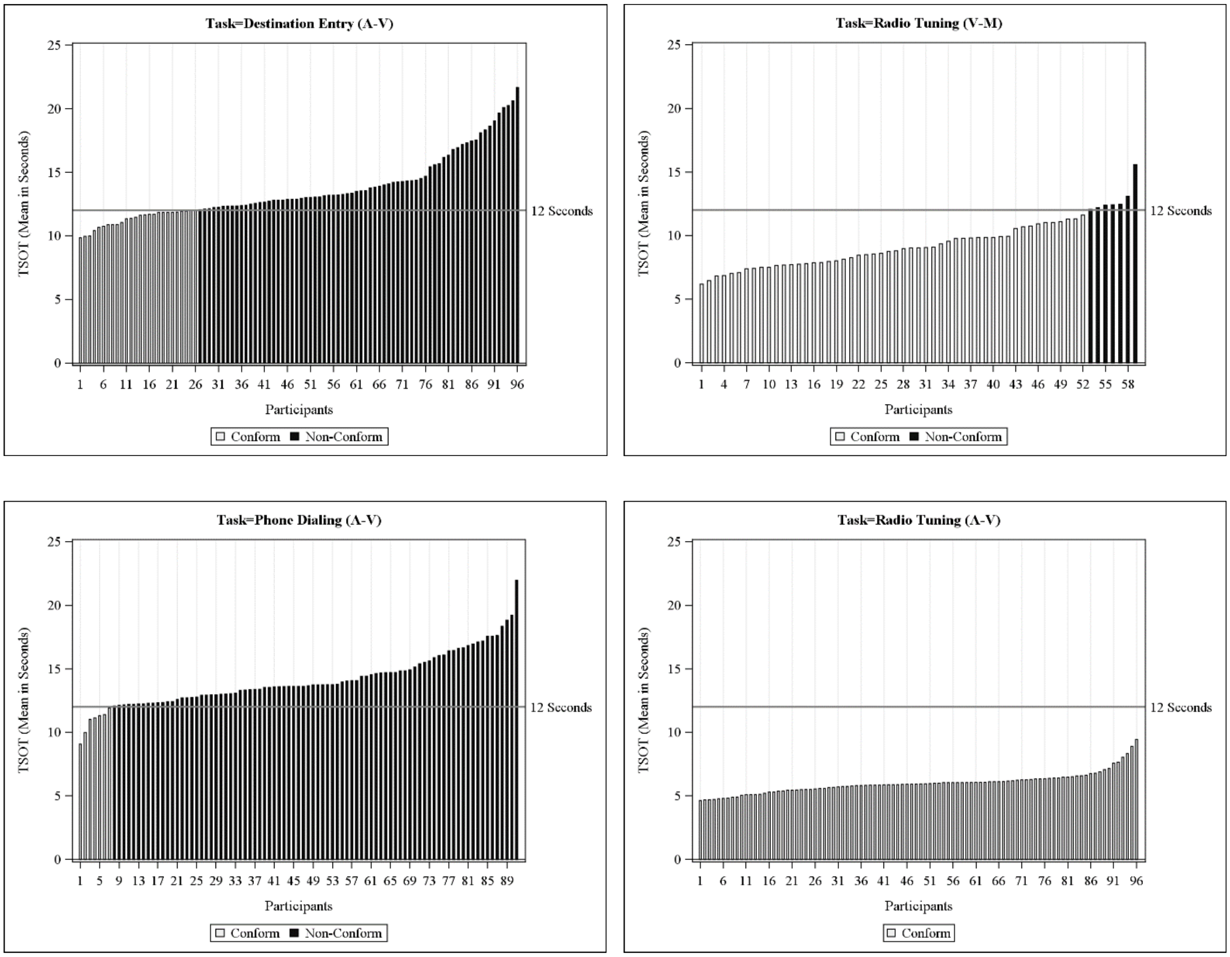


Figure 67. Guidelines TSOT Task Compliance by Task

3.3.3.2 Conformance by Task and Age Group

Figure 68-Figure 71 show the distributions of mean TSOT separated by age group and task, and TSOT above and below 12 s. Age effects are indicated by decreasing percentages of participants who were in conformance with the Guidelines 12-s TSOT criterion with increasing age. The percentages of participants who were in conformance in all the tasks were: 60.2 percent in Age Group 1 (ages 18-24, *N* = 88), 54.1 percent in Age Group 2 (ages 25-39, *N* = 85), 51.1 percent in Age Group 3 (ages 40-54, *N* = 88), and 45.7 percent in Age Group 4 (ages 55+, *N* = 81). Across all the tasks, younger participants (i.e., ages 18-24) were more frequently in conformance than the other age groups.

For the destination entry task, the participant percentages who were in conformance were: 37.5 percent for ages 18-24, 25 percent for ages 25-39, 29.17 percent for ages 40-54, and 16.67 percent for ages 55 and up. The participant percentages who were in conformance for the V-M radio tuning task were: 93.75 percent for ages 18-24 (*N* = 16), 100 percent for ages 25-39 (*N* = 16), 81.25 percent for ages 40-54 (*N* = 16), and 72.73 percent for ages 55 and up (*N* = 11). For the phone dialing task, the participant percentages that were in conformance were: 20.83 percent for ages 18-24 (*N* = 24), 0 percent for ages 25-39 (*N* = 21), 4.17 percent for ages 40-54 (*N* = 24), and 4.55 percent for ages 55 and up (*N* = 22). For A-V radio tuning, 100 percent of the participants in all age groups were in conformance with the 12-s TSOT reference time.

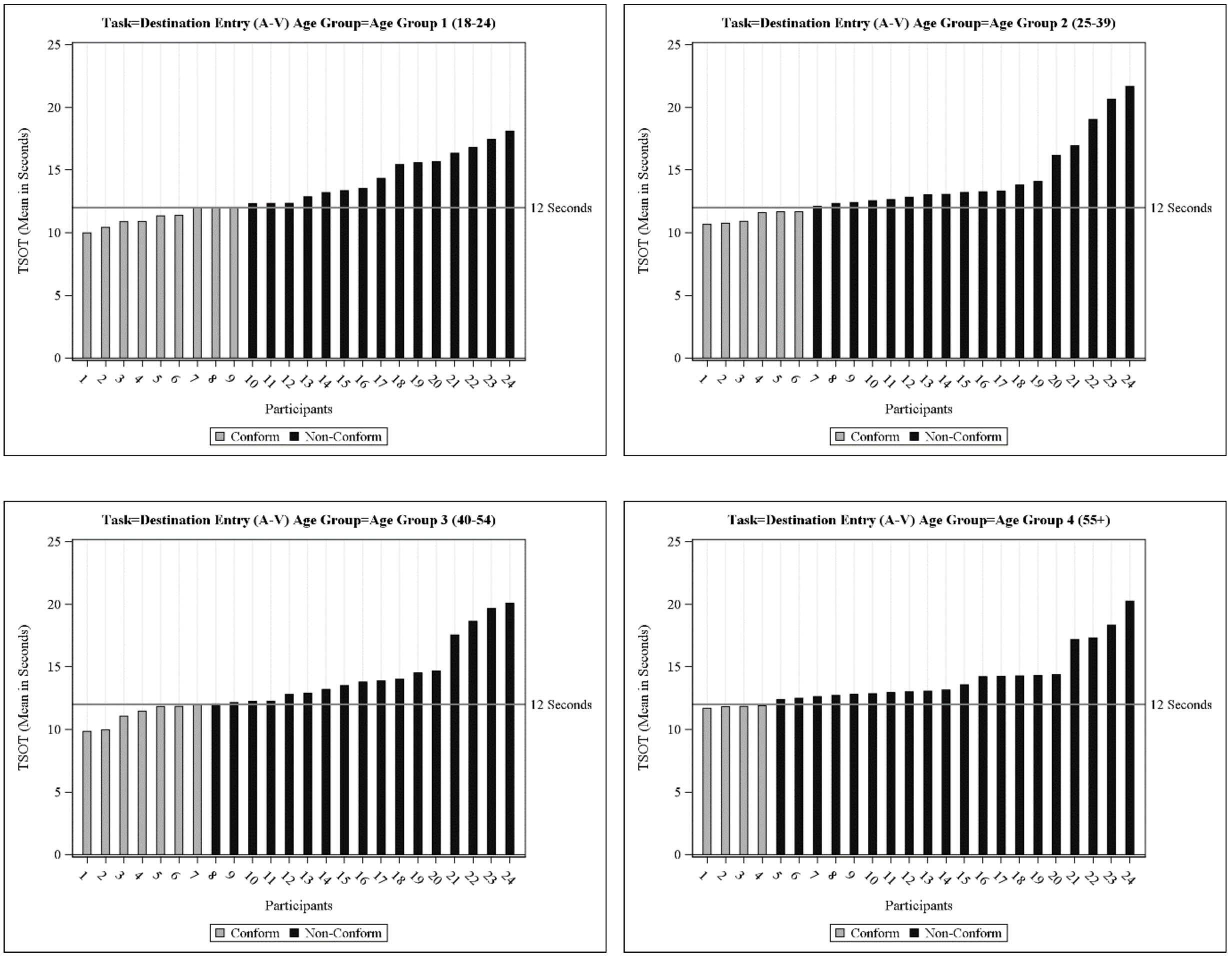


Figure 68. Guidelines TSOT Destination Entry Task Compliance by Age Group

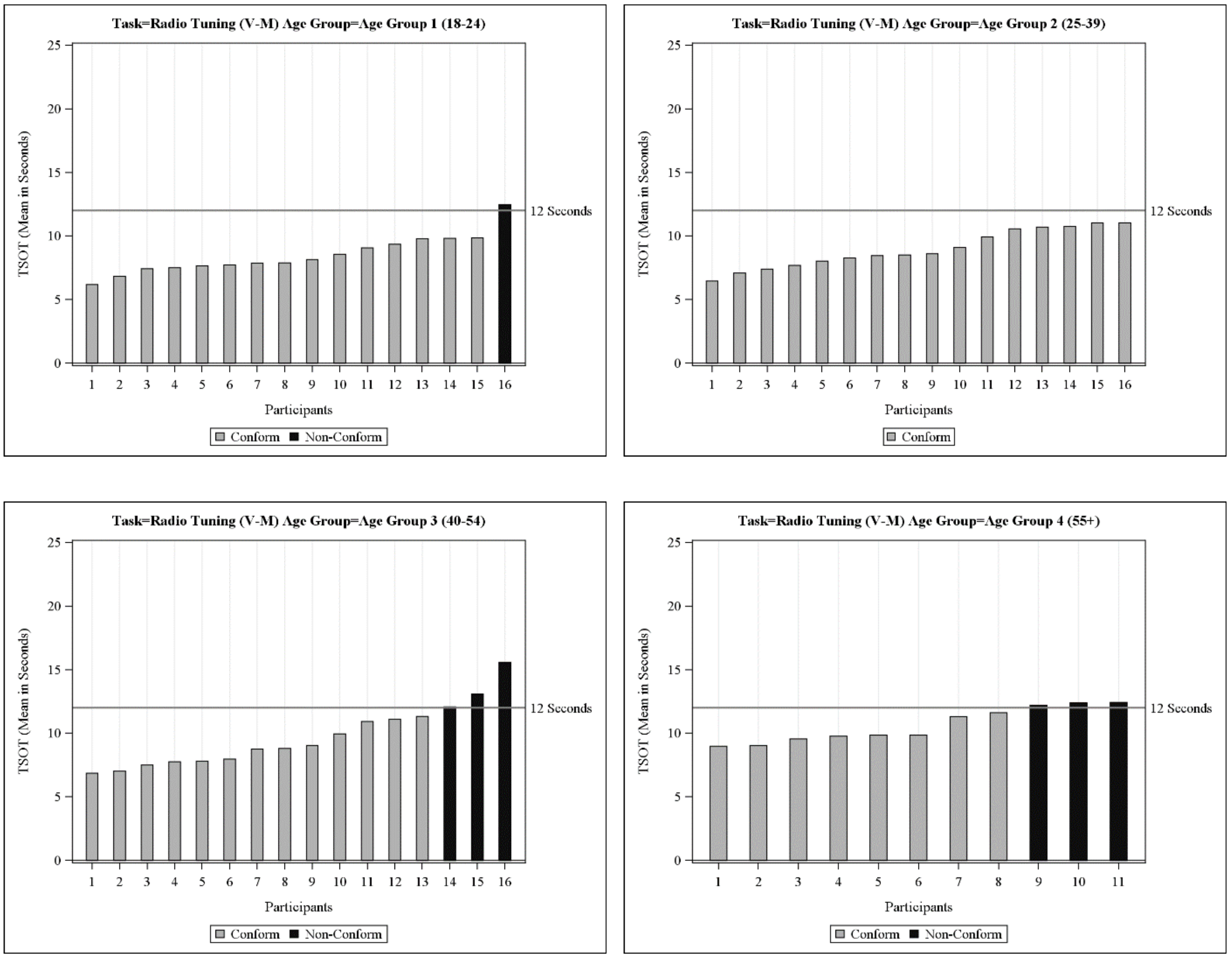


Figure 69. Guidelines TSOT V-M Radio Tuning Task Compliance by Age Group

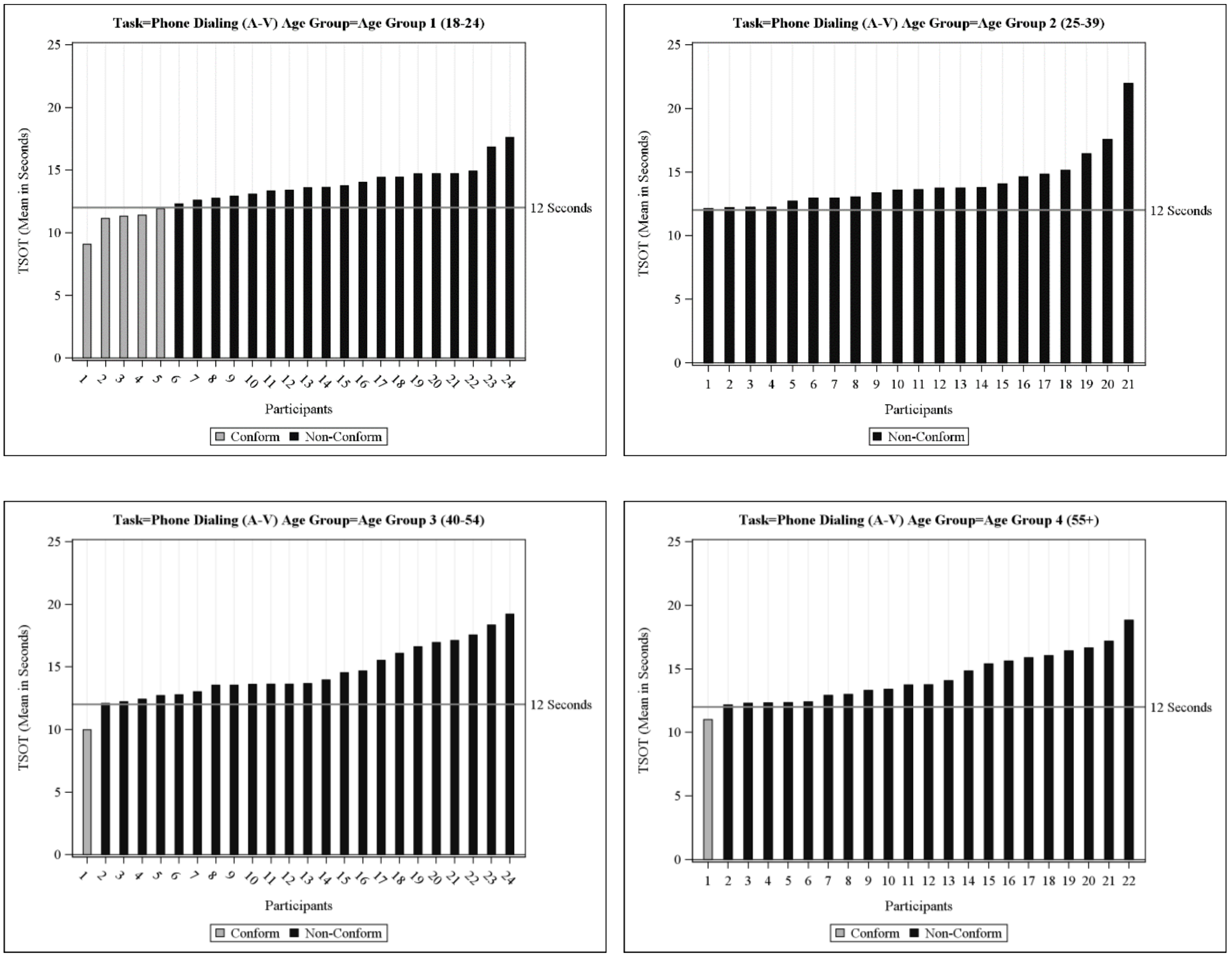


Figure 70. Guidelines TSOT Phone Dialing Task Compliance by Age Group

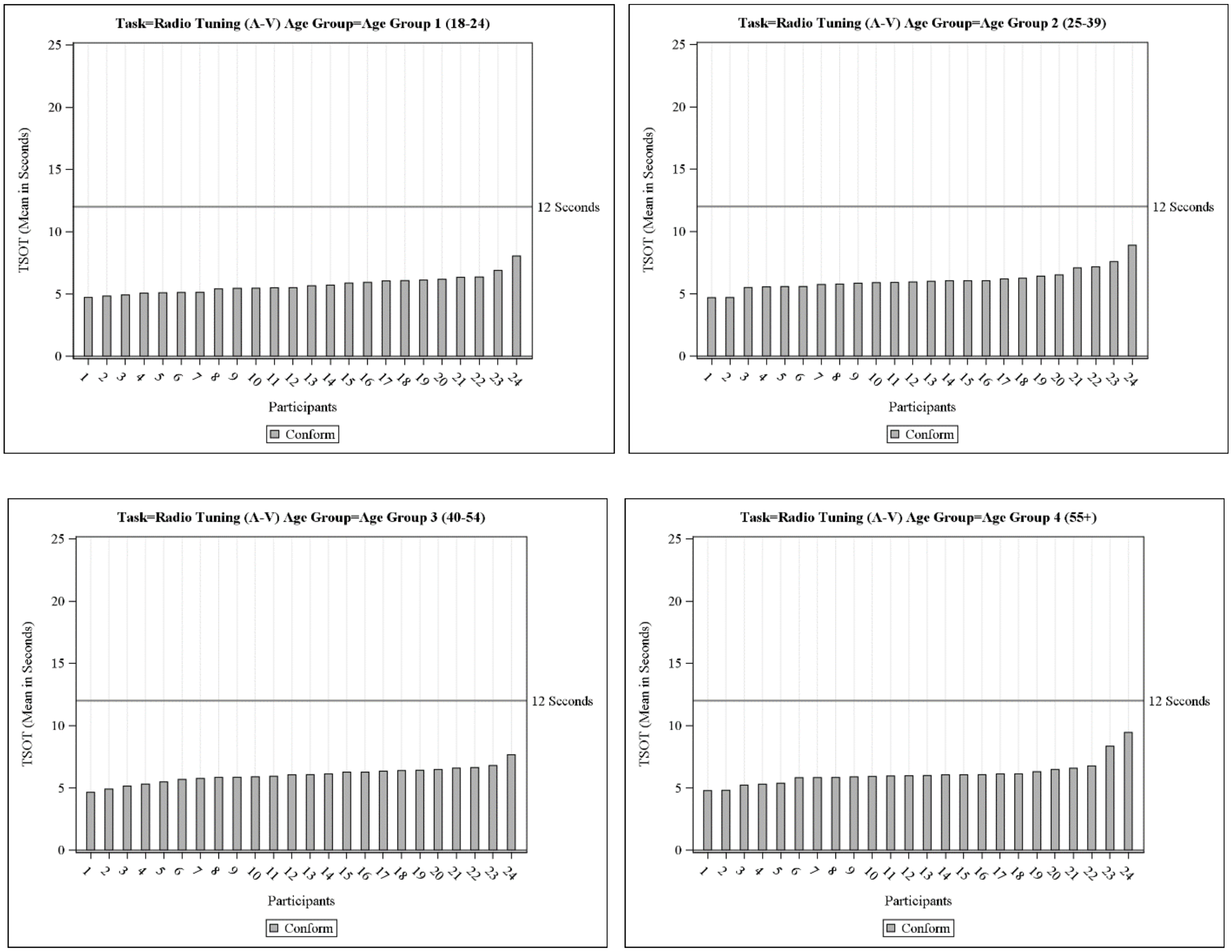


Figure 71. Guidelines TSOT A-V Radio Tuning Task Compliance by Age Group

3.3.3.3 Conformance by Task and Guidelines Sample

Figure 72 shows the mean TSOT for destination entry trials for each of the four 24-person Guidelines samples and the number of participants with conformance or non-conformance outcomes per the 12-s TSOT criterion per task. The NHTSA Driver Distraction Guidelines test outcomes were consistent over these Guidelines groups; all groups had non-conformance outcomes. Guidelines Group 1 (37.5%) had a slightly higher percentage of participants who were in conformance than the other groups (Group 2: 25%; Group 3: 20.8%; Group 4: 25%).

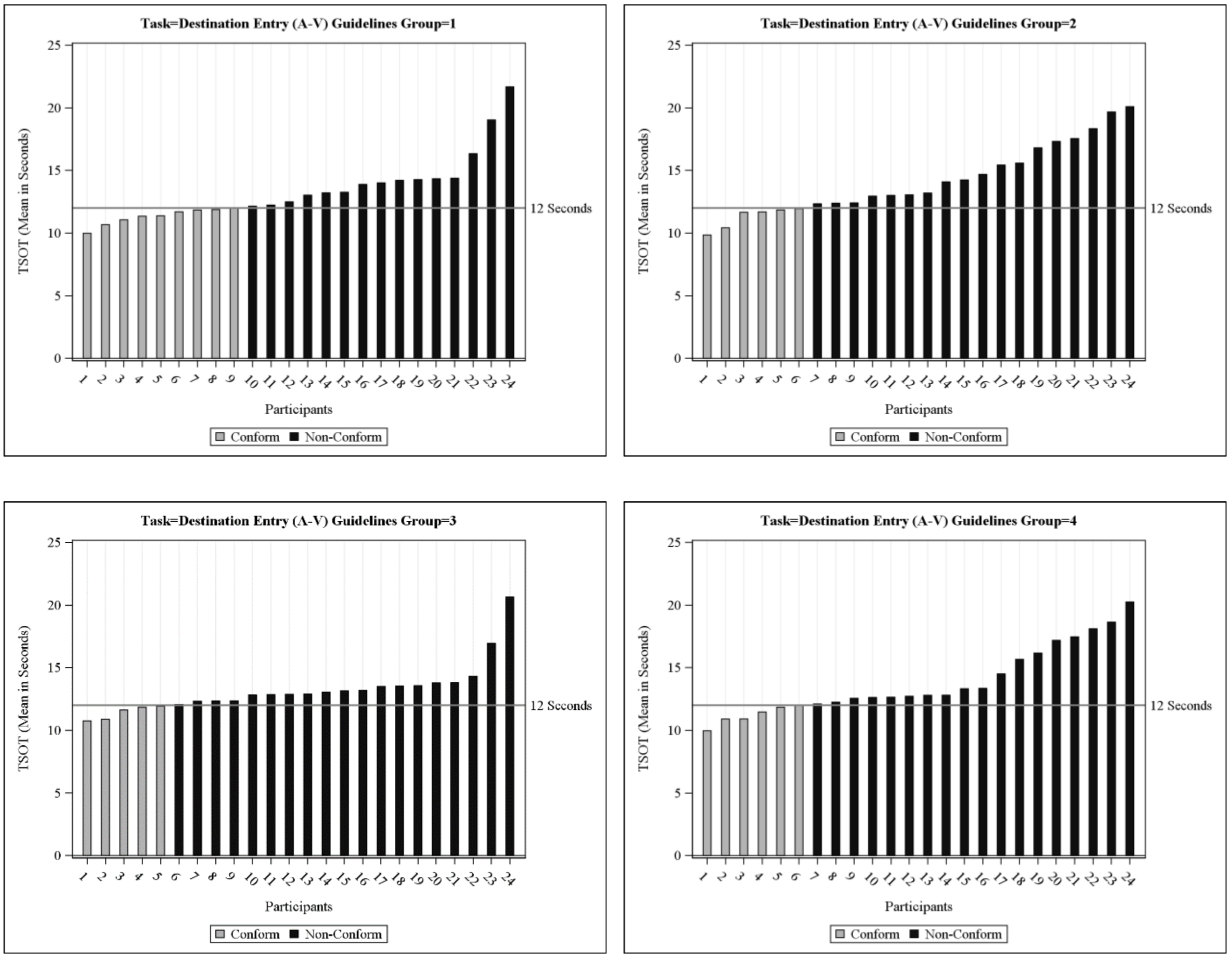


Figure 72. Destination Entry Task Compliance by Task and Guidelines Sample

The mean TSOT during V-M radio tuning for participants in each of the four Guidelines samples (participants with fewer than three error-free trials) and the number of participants who satisfied the TSOT criterion are presented in Figure 73. Conformance percentages were: 84.2 percent of Guidelines Group 1 (N = 19), 93.3 percent of Guidelines Group 2 (N = 15), 84.6 percent of Guidelines Group 3 (N = 13), and 91.7 percent of Guidelines Group 4 (N = 12). Regarding the NHTSA Driver Distraction Guidelines criterion percentage (85%), these groups were not consistent for this task, however, the high number of participants eliminated due to errors precludes formal Guidelines testing.

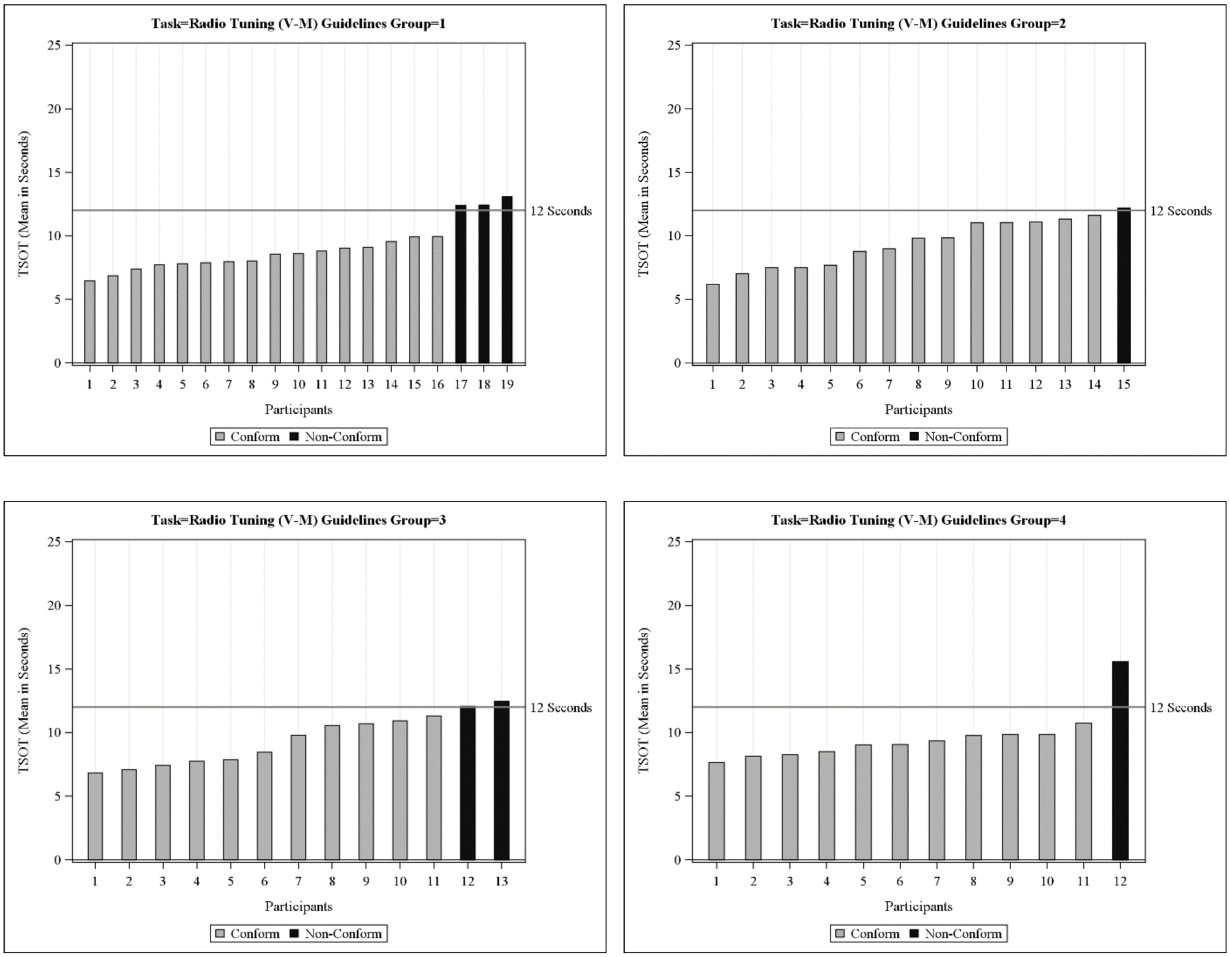


Figure 73. V-M Radio Tuning Task Compliance by Task and Guidelines Sample

Figure 74 shows the mean TSOT for phone dialing trials for each of the four Guidelines samples (not including error trials) and the number of participants who were or were not in conformance with the TSOT criterion. The percentages of participants who were in conformance were: 8.7 percent for Guidelines Group 1 (*N* = 23), 4.6 percent for Guidelines Group 2 (*N* = 22), 8.3 percent for Guidelines Group 3 (*N* = 24), and 9.1 percent of Guidelines Group 4 participants (*N* = 22). When performing the phone dialing task, most participants were not in conformance with the 12-s TSOT reference time.

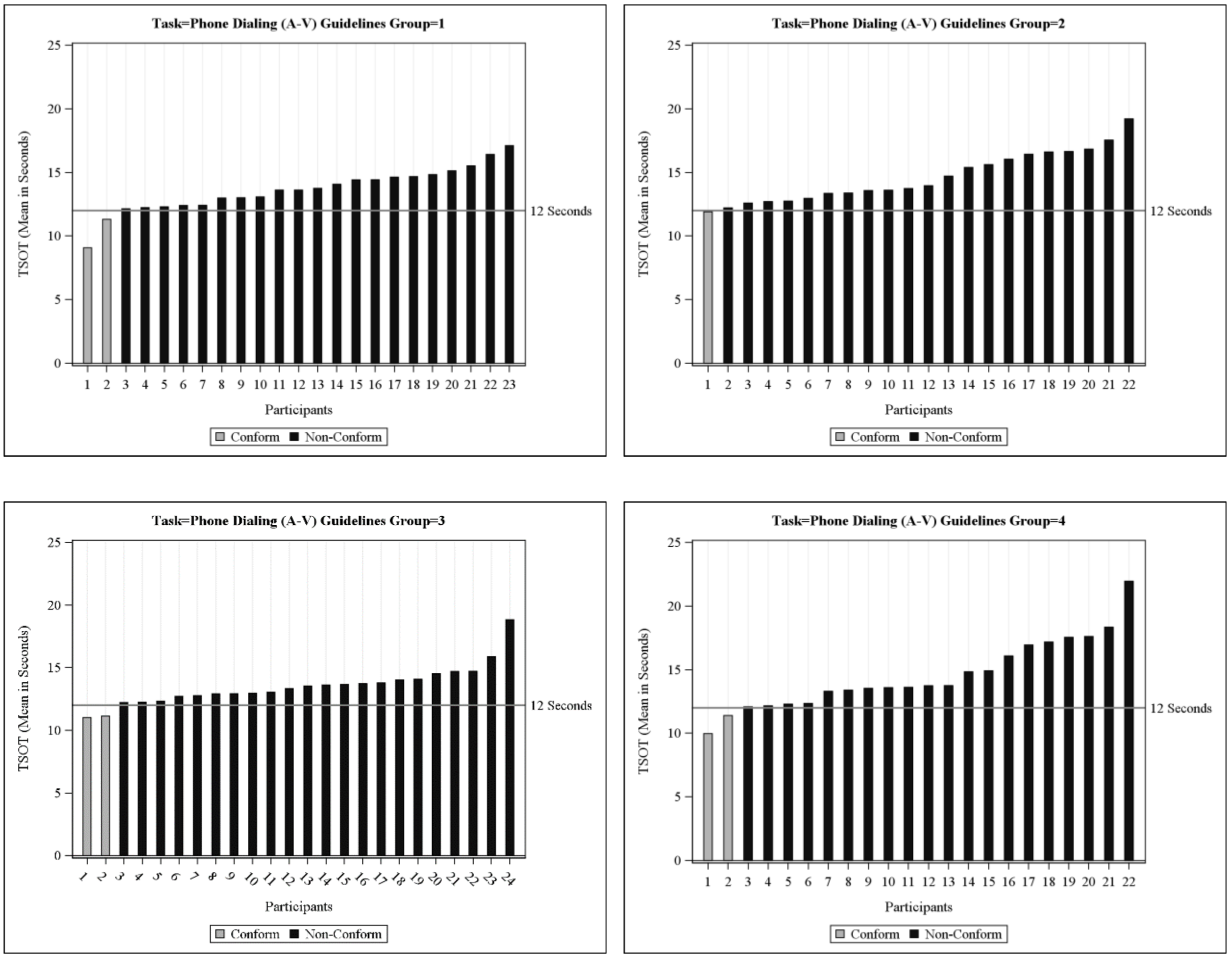


Figure 74. Phone Dialing Task Compliance by Task and Guidelines Sample

Figure 75 displays the mean TSOTs for A-V radio tuning trials for each of the four 24-person Guidelines samples and the number of participants who satisfied the 12-s TSOT criterion. This task is consistent over the NHTSA Driver Distraction Guidelines groups, with all the participants in conformance with the 12-s TSOT reference time.

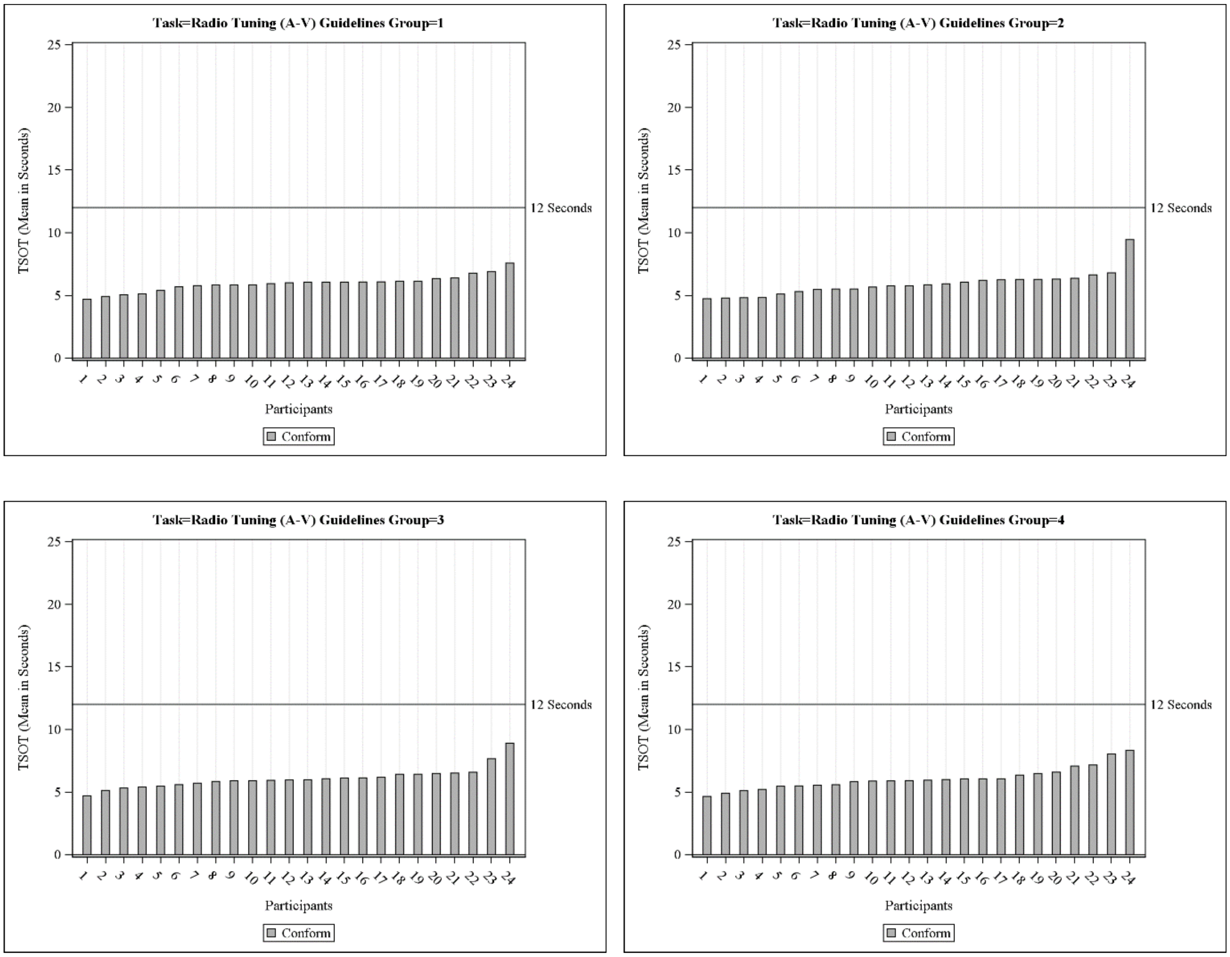


Figure 75. A-V Radio Tuning Task Compliance by Task and Guidelines Sample

3.3.4 Summary of Occlusion Findings

1. TSOT values can differ depending on the way they are computed. TSOT values computed directly from end-of-task button presses resulted in TSOT values that were on average 0.5 s less than those based on the number of shutter-open intervals. The longer TSOT values associated with open-interval computation derive from the inclusion of full open intervals with part of the last interval occurring after the task was completed on some trials. These differences may have no effect on relative assessments of tasks; however, they could affect test outcomes for borderline tests based on absolute TSOT values. Precisely defining the task-completion time is difficult; after task completion, additional time is added by the requirement of participants to report that the task is completed and for the experimenter to press a button to mark the completion time. Having just completed the task, some participants may not fully appreciate the need for a

timely utterance to indicate task completion. Similarly, vigilance among experimenters is required to minimize unwanted delay.

2. Occlusion results indicate that the destination entry and phone dialing tasks used in this experiment generally required more than 12 s of TSOT. The V-M radio tuning and A-V radio tuning tasks generally required less than 12 s of TSOT. Occlusion results for the four tasks were generally consistent across the four 24-person Guidelines samples. The finding that the voice command phone dialing task required more than 12 s of TSOT was unexpected and suggests a design problem with this task.
3. The V-M radio tuning task used in this study demonstrates a potential issue for Guidelines testing. The test outcomes show that the task was close to conformance when data from participants with three or more errors were removed. However, the fact that 39 percent of the participants (37/96) had more than three error trials indicates a significant problem with this task.

3.4 Eye-Glance Analyses

This section presents results pertaining to the third and eighth study objectives.

One objective of the experiment was to use eye-glance data to evaluate in-vehicle tasks relative to NHTSA Driver Distraction Guidelines criteria based on glance durations and TEORT. Figure 76 shows the Areas of Interest (AOI) defined to facilitate analysis of eye-glance data.

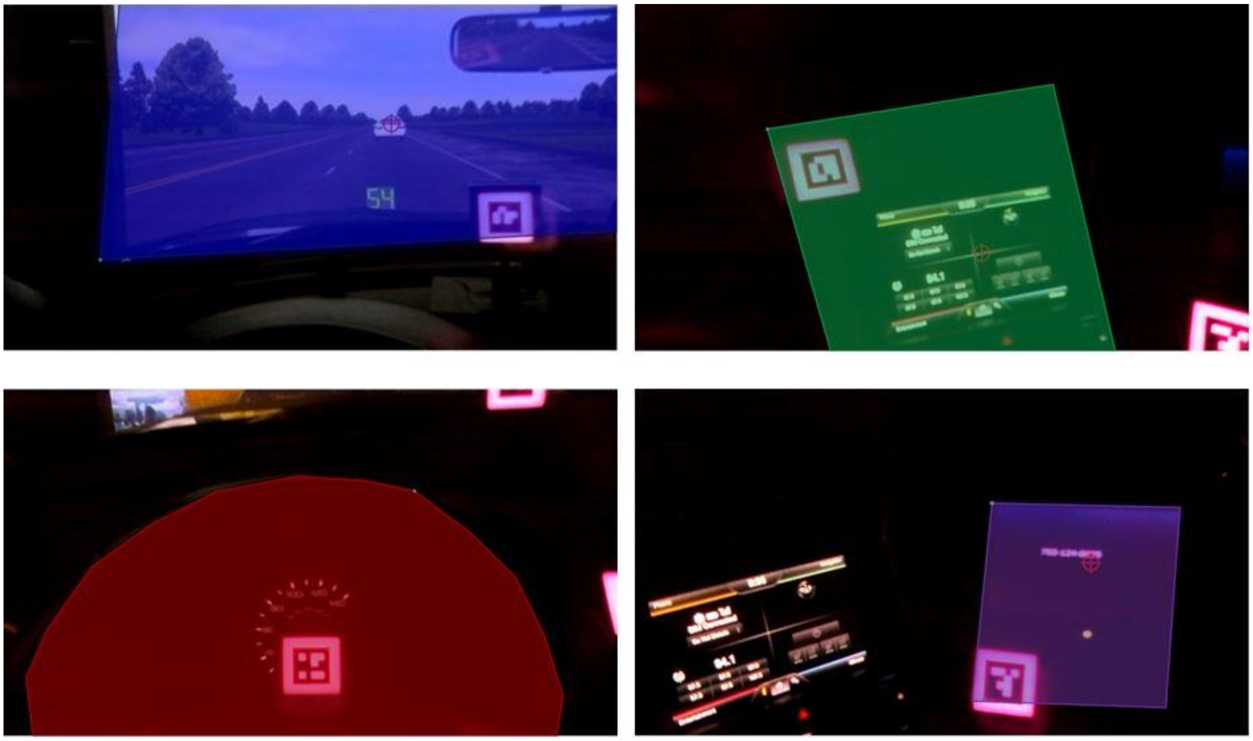


Figure 76. Eye Tracker Areas of Interest

3.4.1 Total Eyes-Off-Road Time by Task

Ninety-six participants assigned to the driving simulator venue performed four in-vehicle tasks while wearing a head-mounted eye tracker. Five participants did not have valid eye-tracker data and were not included in the analyses. Blinks (less than 300 ms) and fly-throughs (less than 120 ms) were removed and gaps filled with D-Lab's glance analysis software. Fly-throughs occurred when the participant glanced past an AOI, not directly at it. However, not all blinks, fly-throughs, and gaps were detected by D-Lab's software, and therefore were not removed.

Gaps in the eye-glance data occurred when the eye-tracker could not determine where the participant was looking. These gaps were assumed to be off-road glances because the road was the largest area of interest and most consistently detected. Anything that was not a road glance was considered an off-road glance. The total duration of glances to the non-road areas of interest (radio console, steering wheel, and task screen) was combined with the total duration of gaps to compute the TEORT for each task.

Gaps between road glances represented the largest proportion of gaps (76.92%). To confirm that most of these gaps were off-road glances, 182 of these gaps (two from each participant) were examined using glance video data. The video data analysis of these gaps indicated that 81.87 percent were single off-road glances, 10.99 percent were road glances, 4.40 percent were blinks, and 2.75 percent consisted of several glances away from the road and back to the road. Thus, our assumption that this category of gaps represents off-road glances is correct for approximately 85 percent of these gaps.

Figure 77 shows the distributions of TEORT for each task. Each value corresponds to one task instance (e.g., a radio station or phone number). The 12-second Guidelines criterion is indicated by a reference line.

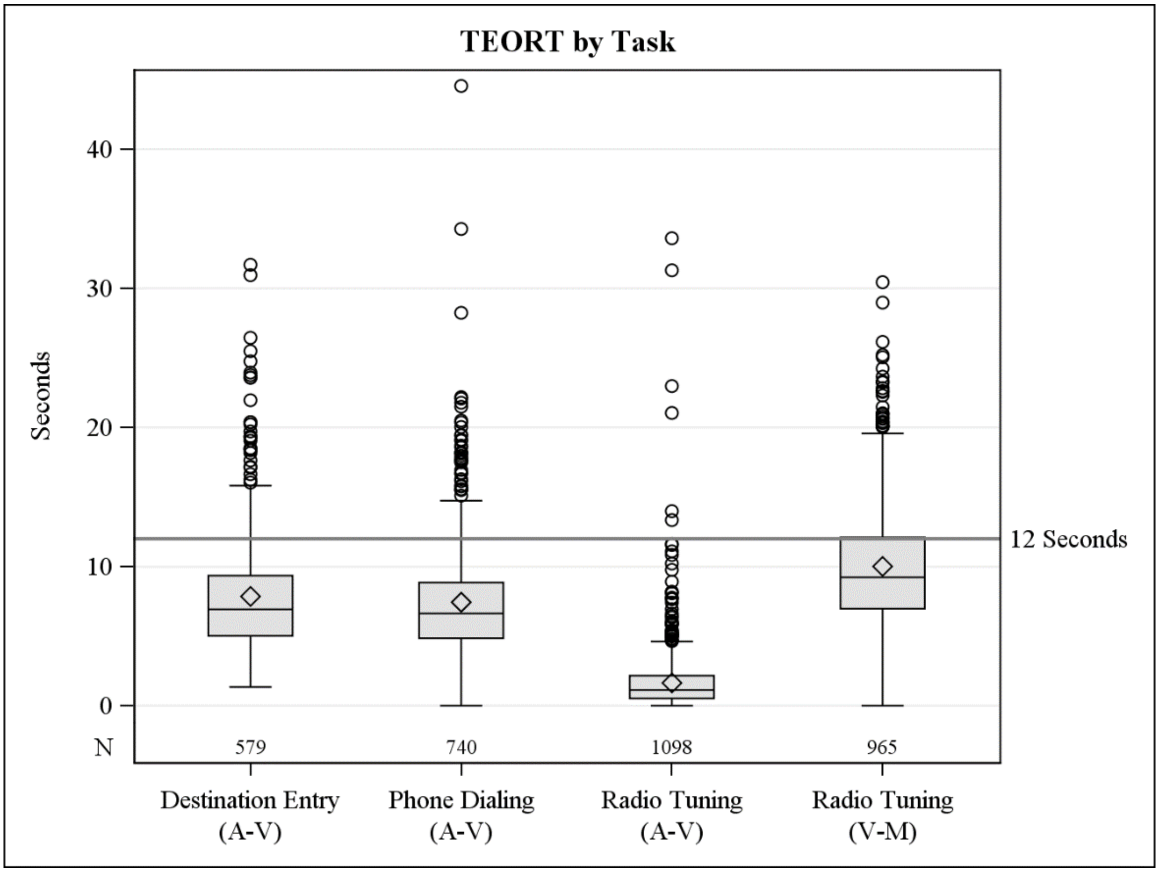


Figure 77. TEORT Distributions by Task

When trials were considered in the aggregate, the 75th percentile TEORT values (upper edge of the boxes) all appeared to be at or below the NHTSA 12-second criterion. V-M radio tuning had the greatest TEORTs (*M* = 10.01, *SD* = 4.16) followed by the A-V tasks (Destination Entry: *M* = 7.85, *SD* = 4.29; Phone Dialing: *M* = 7.44, *SD* = 3.99; A-V Radio Tuning: *M* = 1.63, *SD* = 2.27).

Figure 78 shows the mean TEORT distributions by task, with one value calculated for each task for each participant. Unlike the Guidelines protocol, which specifies that TSOT is computed based on a single instance of each task, the mean TSOT values presented here were based on multiple instances of each task, with more instances of the shorter tasks (e.g., A-V radio tuning) and fewer instances of the longer tasks (e.g., destination entry). Based on the distributions, all tasks except V-M radio tuning had more than 87.5 percent of participants with TEORT values less than the 12-s reference time, which would likely lead to conformance outcomes in NHTSA Driver Distraction Guidelines testing. V-M radio tuning is borderline acceptable, with approximately 80 percent having TEORT values under 12 seconds.

Like the results shown in Figure 77, V-M radio tuning had the highest mean TEORT (*M* = 10.01, *SD* = 2.35), while A-V radio tuning had the lowest mean TEORT (*M* = 1.64, *SD* = 1.21). Destination Entry and phone dialing TEORT means were similar (destination entry: *M* = 7.82, *SD* = 2.63; Phone Dialing: *M* = 7.32, *SD* = 2.64). This is consistent with the expectation that V-M tasks require more off-road glance time than A-V tasks.

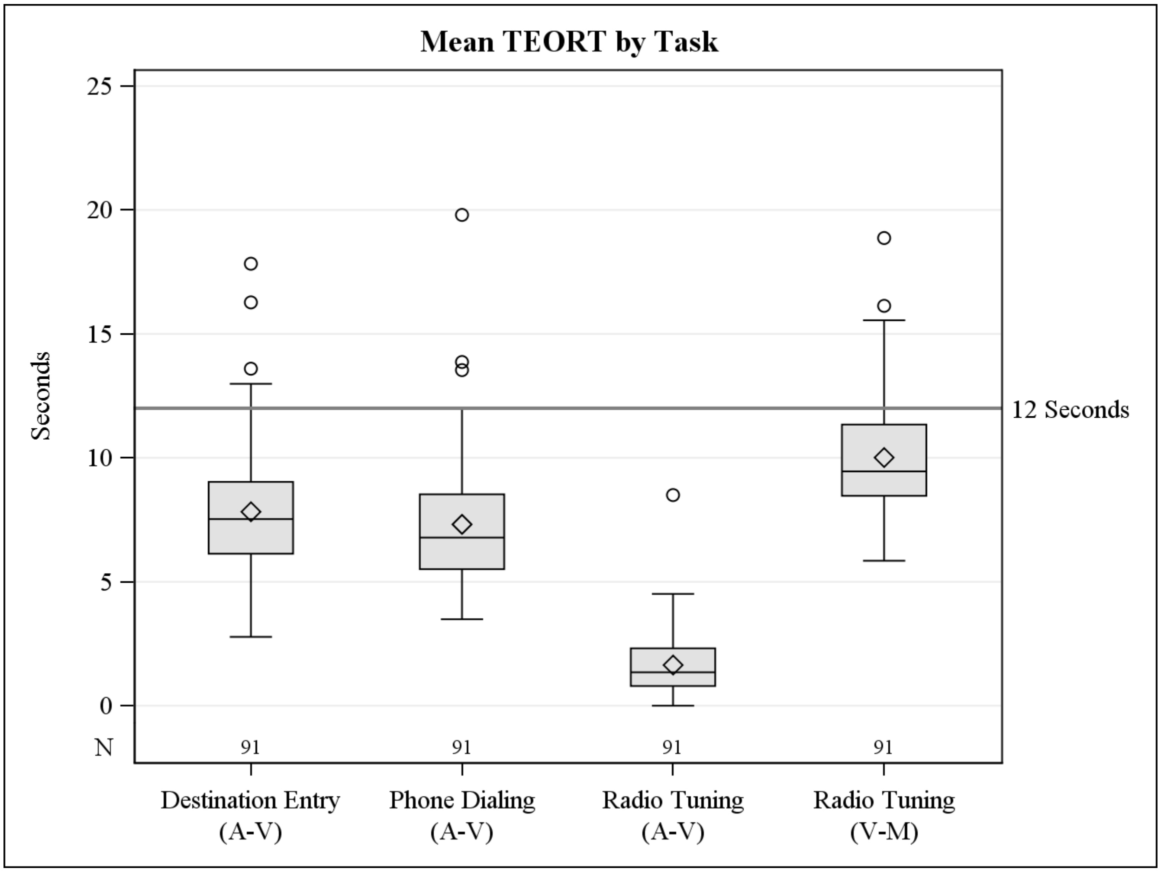


Figure 78. Mean TEORT Distributions by Task

3.4.2 Lead-Vehicle Speed Signal and Eye Glances

This experiment used two lead-vehicle speed conditions to assess the hypothesis that increasing driving task demands would lead to shorter off-road glances and thus, reduced TEORT associated with task completion. The speed conditions included the “Constant” lead-vehicle speed condition specified in the NHTSA Driver Distraction Guidelines, which required minimal attention and a higher-demand “complex” lead-vehicle speed condition, which required drivers to adapt to a constantly-changing speed signal. The complex condition also used a shorter following distance and real-time auditory feedback when the following distance exceeded a pre-defined threshold.

Figure 79 displays the distributions of TEORT for each car-following speed signal. Each value represents one task instance. The overall mean TEORT for the complex car-following trials was smaller and less variable (*M* = 6.07, *SD* = 4.64) than for the constant car-following trials (*M* = 6.63, *SD* = 5.29). Complex car-following trials had proportionately fewer outliers.

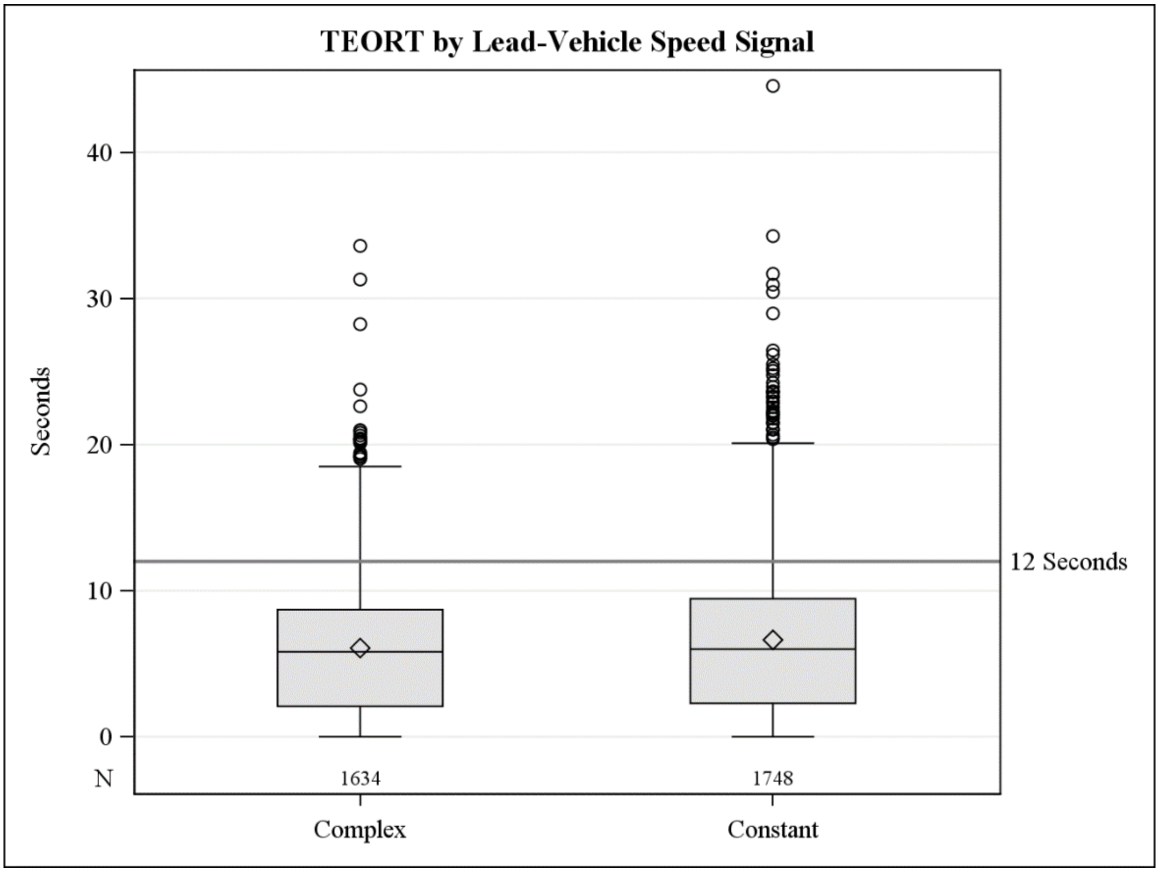


Figure 79. TEORT Distributions by Car-Following Speed Signal: All Tasks

V-M radio tuning was the only task that required relatively continuous off-road glances for task performance. Accordingly, TEORT was examined during this task for the two car-following conditions (Figure 80). The differences seen in Figure 79 became more apparent when the A-V tasks were removed (Complex: *M* = 9.63, *SD* = 3.82; Constant: *M* = 10.37, *SD* = 4.44).

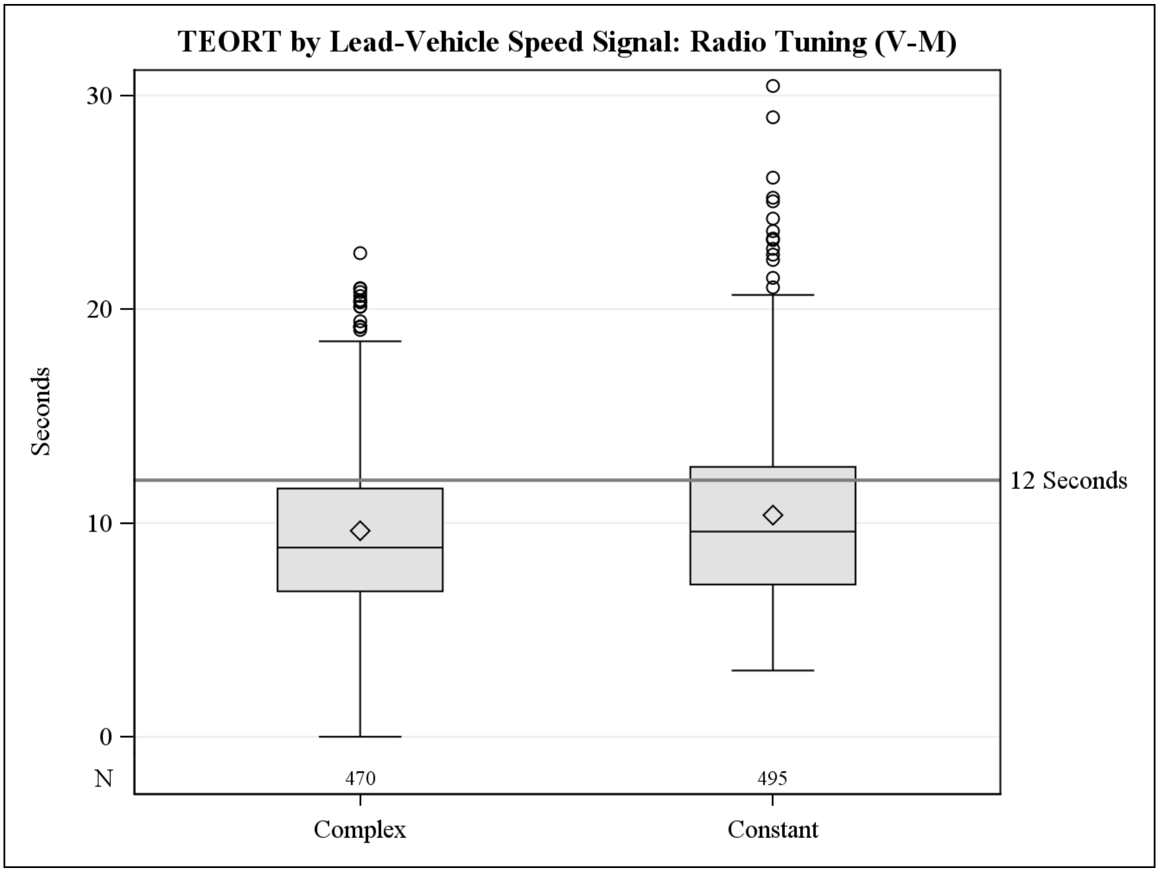


Figure 80. TEORT Distributions by Car-Following Speed Signal: V-M Radio Tuning

The previous presentations have considered data in the aggregate. Next, comparable distributions of mean TSOT values are presented with each point summarizing glance data from all instances of a task for a single participant. Figure 81 shows the mean TEORTs for all task conditions for the two car-following speed conditions.

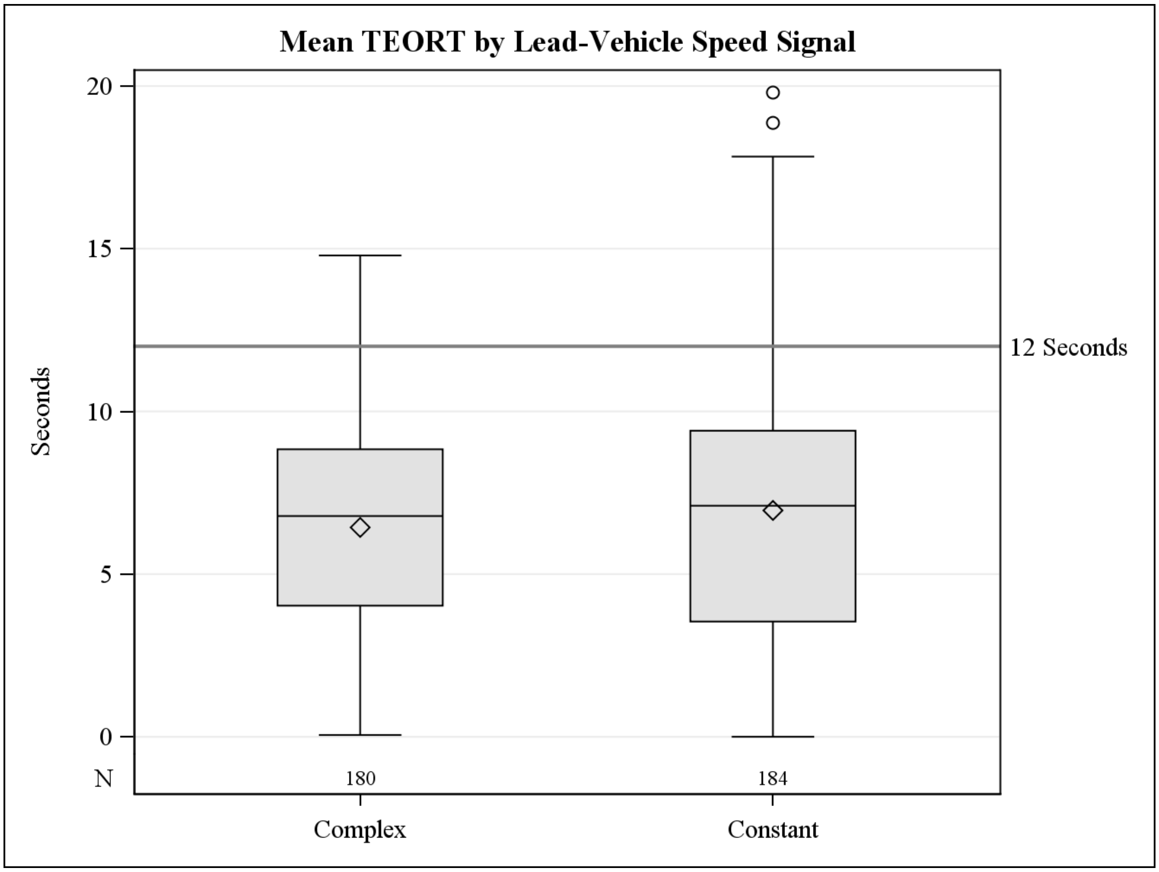


Figure 81. Mean TEORT Distributions by Car-Following Speed Signal: All Tasks

The trend is the same as for the aggregate distributions: participants accumulated less glance time away from the road during Complex car-following (*M* = 6.43, *SD* = 3.51) than during the Constant car-following condition (*M* = 6.96, *SD* = 4.13). The increased variability observed in the aggregated data is also evident in the distributions of mean TEORT values.

The mean TEORT for the two car-following speed conditions during V-M radio tuning is displayed in Figure 82.

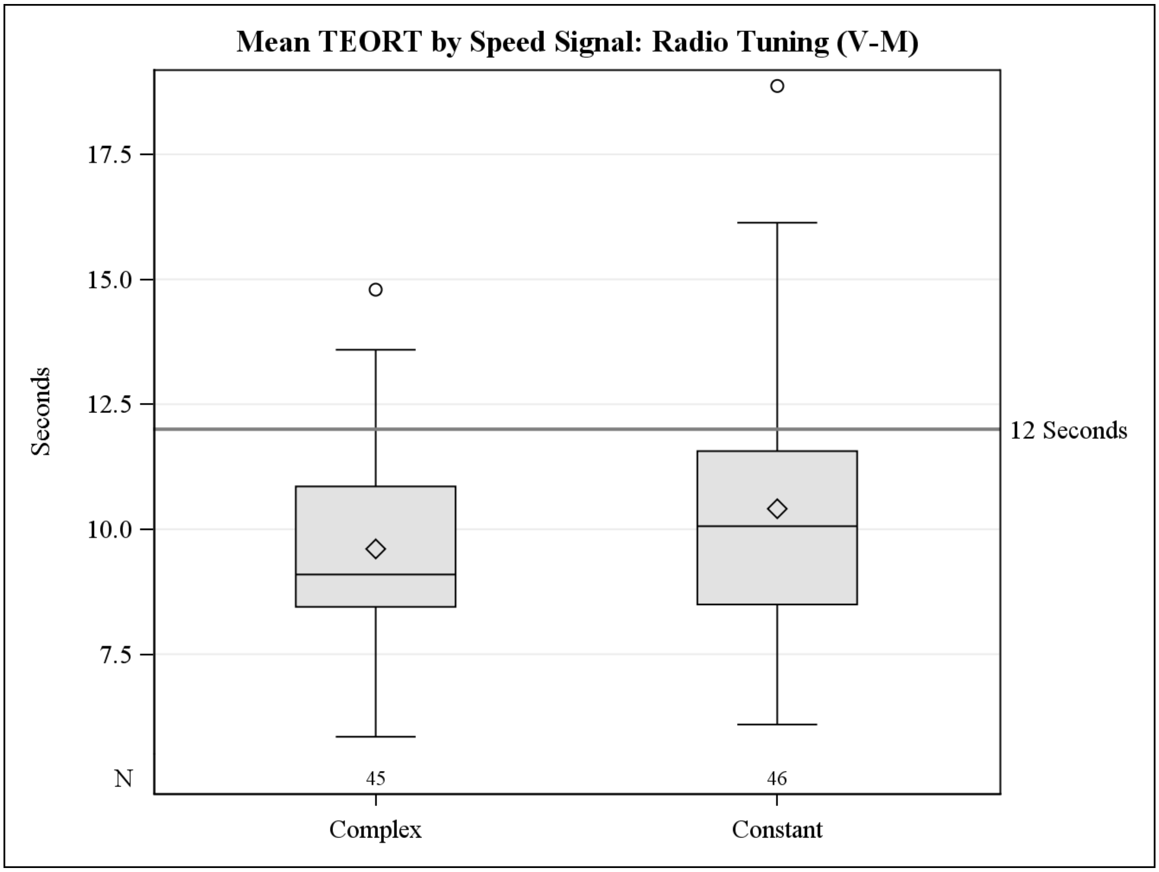


Figure 82. Mean TEORT by Speed Signal: V-M Radio Tuning

These differences are consistent with those in the aggregated data set: Complex car following during V-M radio tuning was associated with shorter mean TEORTs ($M = 9.61$, $SD = 1.99$) than Constant car following ($M = 10.41$, $SD = 2.61$). The complex car-following condition was also associated with reduced variability among participants, as shown above.

Statistical testing was done for this comparison because the data points were constructed so that each point satisfied independence requirements. Visual-Manual Radio Tuning was chosen because it is the only task for which off-road glances are required to perform the task. The difference between means was found to be not statistically different, $t(89) = 0.10$; however, the Cohen's d associated with this difference was $d = 0.34$, which is midway between a small and medium effect. As shown in the figure, the main effects of the increased car-following difficulty appear to be a modest decrease in mean TEORT and a narrowing of the distribution.

We examined distributions for the remaining two NHTSA Driver Distraction Guidelines glance metrics separated by lead-vehicle speed signal. Figure 83 presents distributions of mean glance durations for the respective lead-vehicle speed conditions. Participants typically completed 10 radio tuning task instances in two driving trials. For each participant, the mean duration was computed using off-road glances from all task instances, which was typically between 50 and 150 glances. The distributions have the following frequencies: $N_{\text{complex}} = 45$; $N_{\text{constant}} = 46$, which reflects the numbers of participants with good glance data.

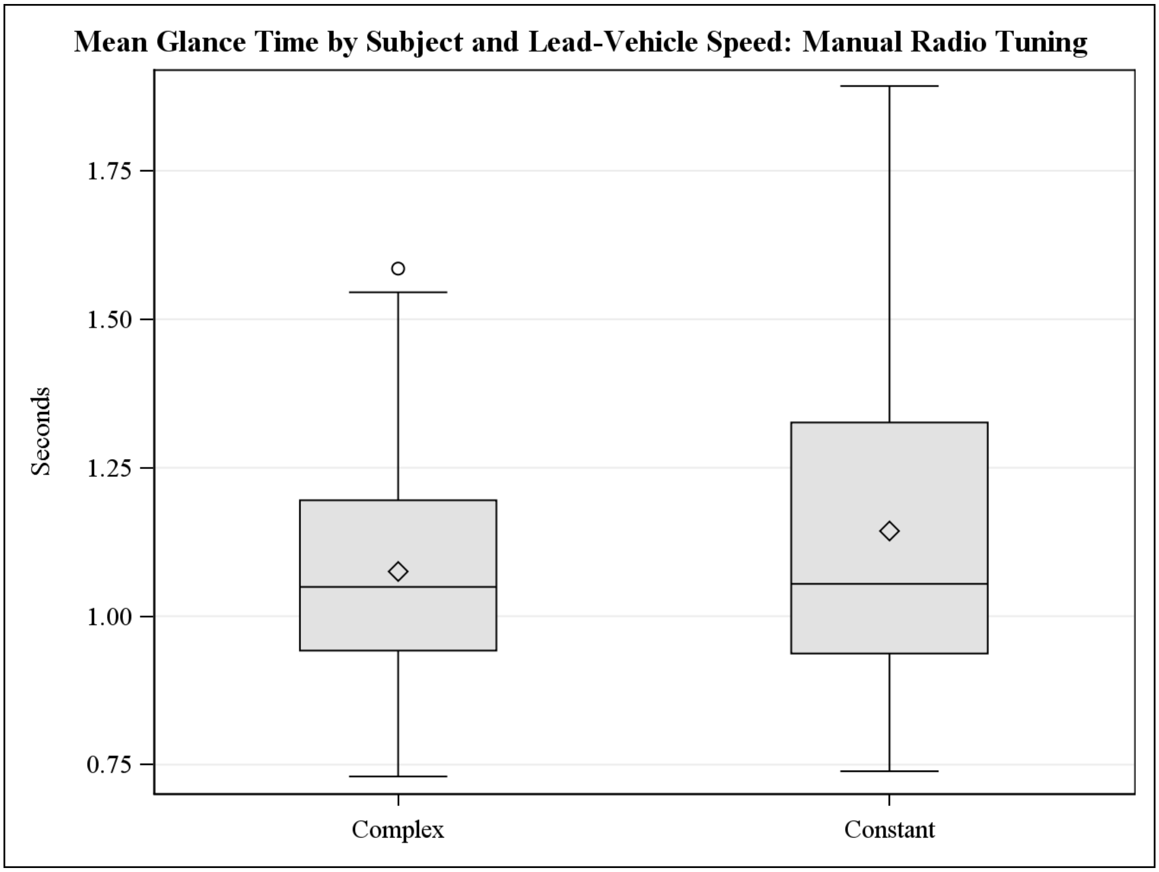


Figure 83. Mean Glance Time by Subject and Lead Vehicle Speed: V-M Radio Tuning

Most evident are differences in the overall and interquartile ranges, both of which indicate reduced variability among drivers in the complex speed condition. Values up to and including the median were generally consistent for the two distributions; differences are most apparent in the upper 50 percent of each distribution. Mean values, indicated by diamonds in the figure, differ (M_{constant} = 1.14 s; M_{complex} = 1.075 s), which indicates a small (6%) reduction in mean glance duration in the complex condition. This difference was not statistically significant, $t(89) = -1.30$, $p = .20$; however, Cohen's $d = 0.27$ indicates a small effect. The corresponding values when computing a non-parametric t-test with ranks of means was: $t(8) = -0.83$, $p = .41$; $M_{\text{complex}} = 43.67$, $SD = 36.25$; $M_{\text{constant}} = 48.28$, $SD = 28.09$, Cohen's $d = 0.17$. The effect size and statistical test outcome were virtually identical when the mean durations were computed using task instance means for each participant.

We also computed the proportion of glances longer than 2.0 s for each participant, first collapsed across all task instances and then by task instance. Figure 84 presents the distributions of proportions collapsed across all task instances for each participant. The distributions have the following frequencies: $N_{\text{complex}} = 45$; $N_{\text{constant}} = 46$.

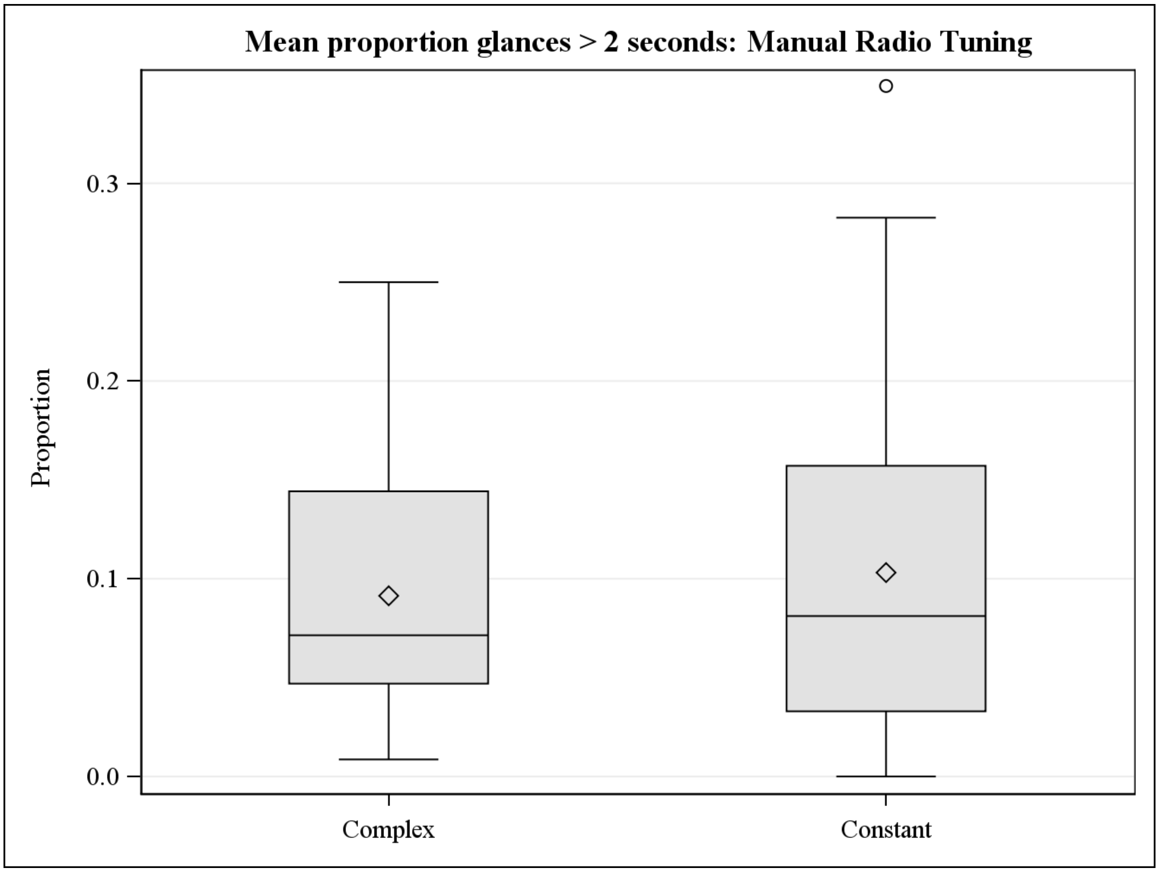


Figure 84. Proportion of Glances Longer Than 2 Seconds: V-M Radio Tuning

As above, the notable features of this comparison are the differences in overall and interquartile ranges of the two distributions. For this metric, shrinkage of variability was evident in both directions. The mean values revealed a small difference (16%), with *M* = 0.091 for the complex condition versus *M* = 0.103 for the constant condition. This difference was not statistically significant, *t*(89) = -0.73, *p* = .47. Cohen's *d* for this effect is 0.16, which falls below the 0.20 threshold for a small effect. The change in distribution range is reflected in the difference between standard deviations (SD_{complex} = 0.07 vs. SD_{constant} = 0.09). Corresponding results of statistical testing using non-parametric testing with ranks of mean values had similar results, *t*(89) = -0.29, *p* = .77; M_{complex} = - 45.18, SD_{complex} = 24.75; M_{constant} = 46.80, SD_{constant} = 28.19. Cohen's *d* for this effect was *d* = 0.06, indicating a negligible effect. These results confirm that differences between lead-vehicle speed conditions are evident primarily in the shape but not the position of the respective distributions.

3.4.3 Effects of Lead-Vehicle Speed Signal by Age Group

The sample of simulator participants covered a wide range of ages. To explore the effects of different lead-vehicle speed signals on different age groups, the respective distributions were separated by age group and speed signal. Figure 85 presents the distributions of mean glance duration by age group and lead-vehicle speed condition.

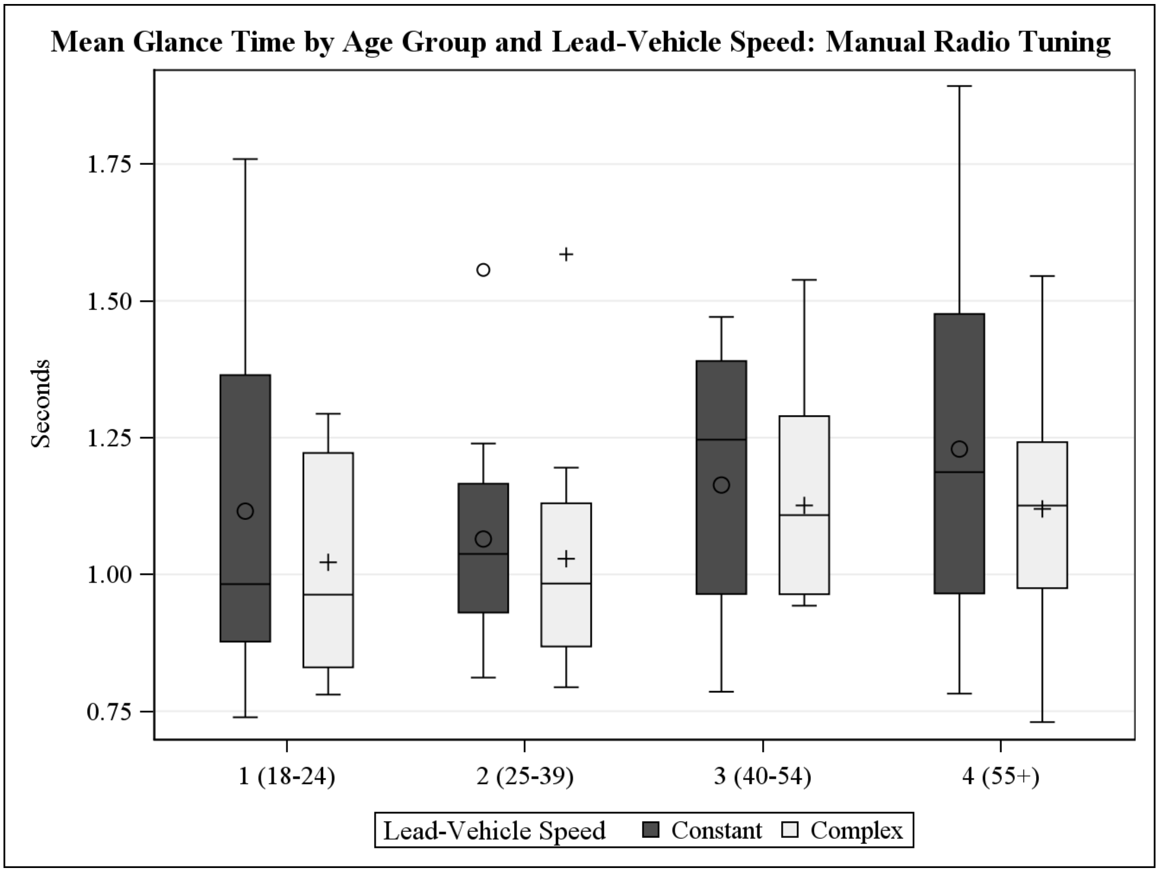


Figure 85. Mean Glance Time by Age Group and Lead Vehicle Speed: V-M Radio Tuning

The pattern of differences evident in the overall distribution was apparent in Age Groups 1, 3, and 4. Specifically, the Constant Lead-Vehicle Speed condition was associated with more variability among the individual mean glance durations and with slightly elevated overall mean values. Increased variability was not apparent for Age Group 2; however, there was a small difference between distribution mean values. Distributions of the proportions of glances longer than 2 seconds during manual radio tuning by age group and Lead-Vehicle Speed Condition are presented in Figure 86.

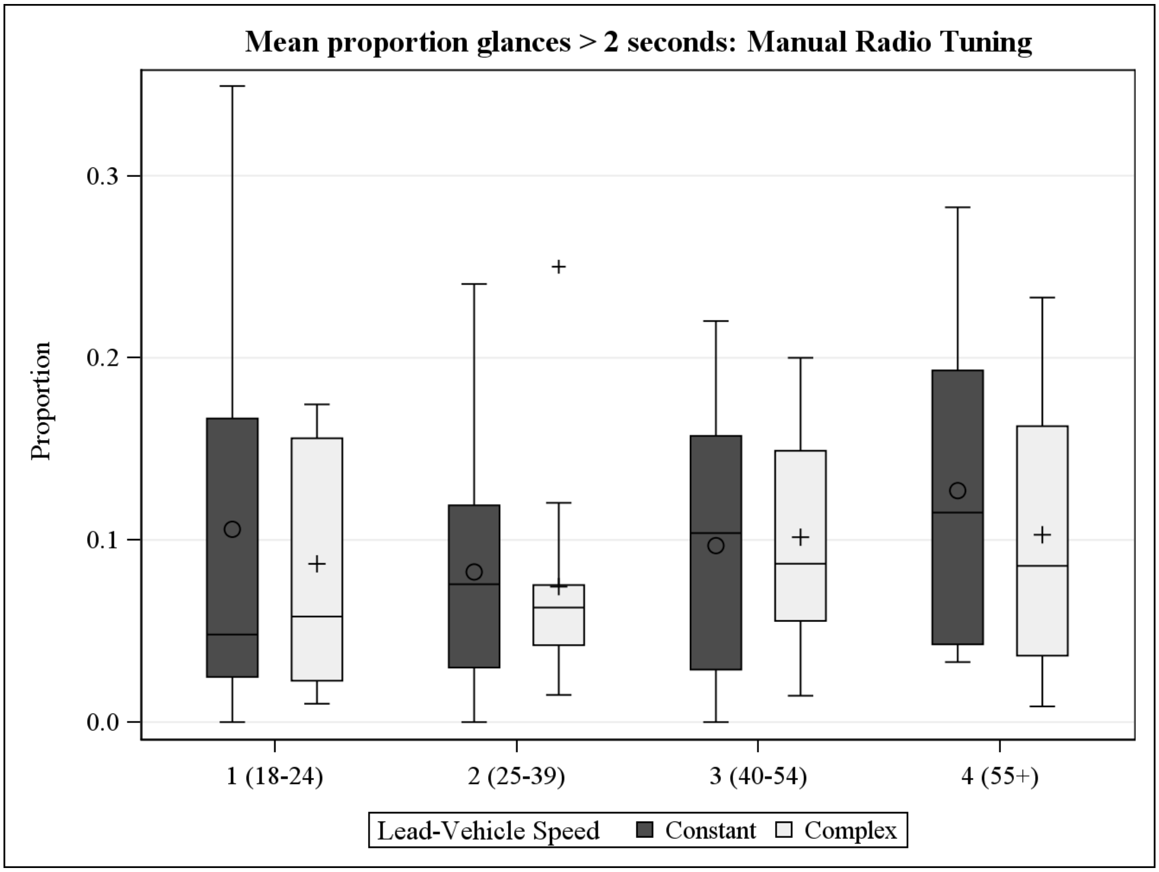


Figure 86. Proportion of Glances Greater Than 2 Seconds by Age Group and Lead-Vehicle Speed Signal: V-M Radio Tuning

For this metric, Age Group 2 reveals the greatest reduction in variability associated with the Complex condition. For Age Group 1, the reduction is primarily evident for the upper 25 percent of the distribution. Distributions of TEORT for each combination of age group and Lead-Vehicle Speed Condition are presented in Figure 87.

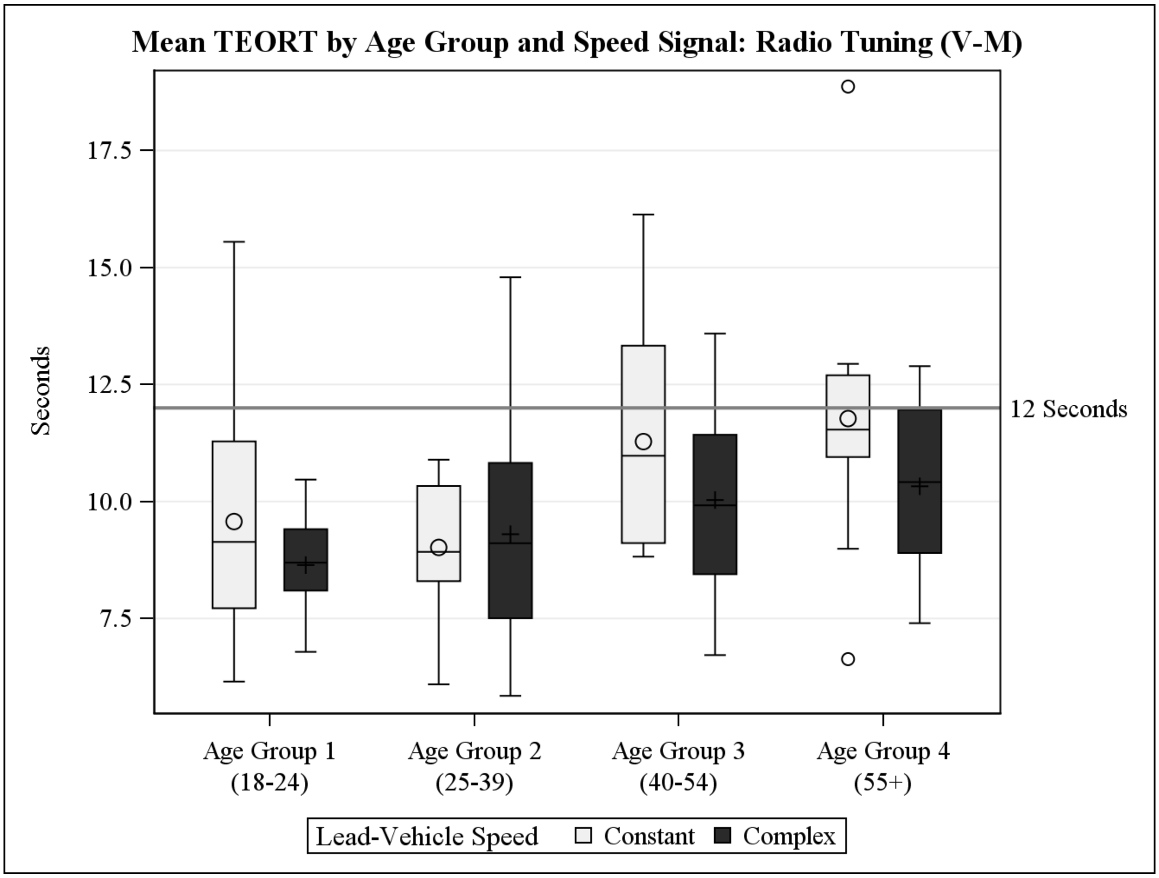


Figure 87. Mean TEORT by Age Group and Lead-Vehicle Speed Signal: V-M Radio Tuning

The effect of lead-vehicle speed was most evident for Age Group 1, which represents the youngest drivers. Mean TEORT values reveal reductions in both the mean and variability of the distribution. For Age Group 3, the effect of the Complex speed signal was to move the entire distribution lower as well as reducing the variability.

3.4.4 Task-Completion Time Analyses

One objective of this study was to determine whether occlusion testing can be used to assess tasks performed using voice commands and auditory feedback. Because occlusion assumes that task completion requires continuous glances to the in-vehicle system, one way to determine the usefulness of occlusion testing is to examine the off-road glance time (TEORT) associated with each task obtained in the simulator in relation to the task-completion time. Distributions of the time it took participants to complete each in-vehicle task are shown in Figure 88. One value was calculated for each task instance.

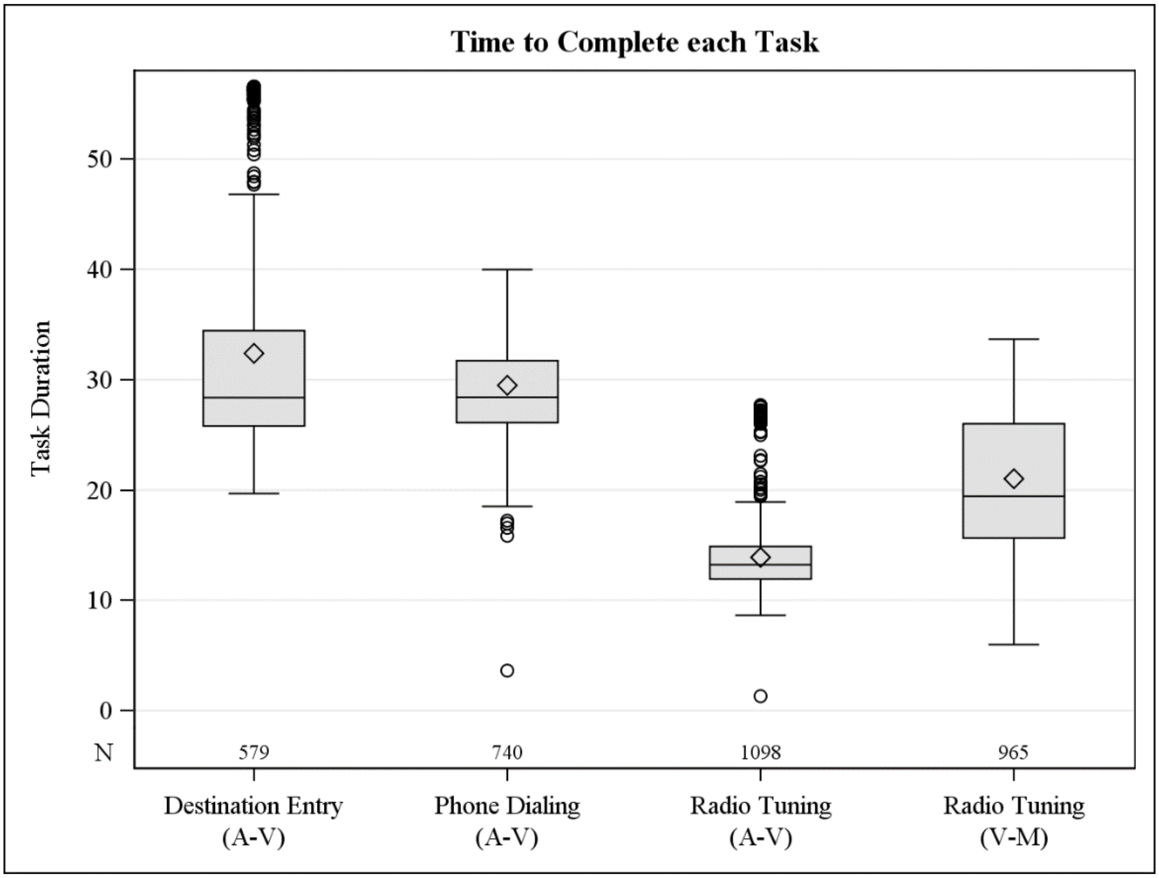


Figure 88. Task Duration Distributions by Task

Participants took the longest amounts of time to complete destination entry (*M* = 32.4, *SD* = 10.30) and Phone Dialing (*M* = 29.49, *SD* = 5.55). Additionally, the destination entry task time was associated with the most variability over task instances, although this appears to be due to outliers. The two radio tuning tasks took less time to complete (V-M radio tuning: *M* = 21.03, *SD* = 7.00; A-V radio tuning: *M* = 13.90, *SD* = 3.28).

We examined correlations between TEORT and task-completion time. Overall, the correlation was $r = .63$, $n = 2282$, $p < .0001$. This correlation indicates that TEORT values were moderately correlated to task-completion times; however, with $r^2 = .40$, the implication is that less than half of the total variation in task-completion time is associated with TEORT. Correlations were examined for each task separately. The results are presented in Table 24.

Table 24. Correlations Between TEORT and Task-Completion Time by Task

Task	**N**	**Correlation (r)**	**p**
Destination Entry	579	.67	< .0001
V-M Radio Tuning	965	.74	< .0001
Phone Dialing	740	.50	< .0001
A-V Radio Tuning	1098	.42	< .0001

The correlation values differed by task. The highest correlation was associated with the V-M Radio Tuning task, which indicates that task-completion time is more consistently related to

TEORT than for other tasks. The lowest two correlations were associated with the simpler voice-based tasks (i.e., phone dialing and A-V radio tuning). Weaker correlations between TEORT and task-completion time may reflect the fact that tasks can be performed effectively with voice commands and thus require a smaller percentage of task-completion time devoted to off-road glances. The weaker correlations associated with the simpler voice-based tasks and the stronger correlation associated with the V-M task (V-M Radio Tuning) are consistent with this interpretation.

The occlusion protocol is structured so that TSOT equals exactly half of task-completion time, thus task-completion time and the viewing time required for task performance are assumed to be perfectly correlated. Accordingly, occlusion is most suitable for assessing tasks with a strong relation between these two metrics and can be expected to provide invalid estimates of required viewing time for tasks that have relatively low correlations between task-completion time and actual on-road viewing time as measured by TEORT.

To examine the relation between TEORT and task-completion time more directly, TEORT was divided by task-completion time for each task instance. Distributions of the resulting proportion of task-completion associated with off-road glances were examined for each task type. The results are summarized in Table 25.

Table 25. Proportion of Task-Completion Time Devoted to TEORT by Task

Task	**N**	**Mean**	**SD**
Destination Entry	579	0.24	0.09
V-M Radio Tuning	965	0.48	0.12
Phone Dialing	740	0.25	0.11
A-V Radio Tuning	1098	0.11	0.14

As with the correlations, V-M radio tuning had the greatest proportion of task-completion time devoted to off-road glances. The results provide additional support for the conclusion that occlusion is best suited for assessing V-M tasks, defined operationally here as those with approximately half of task completion time devoted to off-road glances necessary for task completion. It should be noted however, that even for V-M radio tuning, the range of individual task-completion proportions devoted to off-road glances varied considerably, reflecting the fact that even for V-M tasks, performance differs widely when viewing time is not strictly regulated.

3.4.5 Summary of Eye Glance Findings

1. Overall, V-M radio tuning had a higher TEORT than the A-V tasks. This is consistent with the expectation that V-M tasks require more off-road glances than A-V tasks.
2. The more demanding complex car-following was associated with reductions in variability for three glance metrics, including TEORT, mean glance duration, and proportion of glances longer than 2 seconds, relative to less demanding constant car-following. The mean values for each metric revealed small reductions in the complex condition, however, none of these effects was statistically significant. Together, the results suggest that increasing the car-following demands has the potential to reduce the variability and mean values of the glance metrics.

3.5 Heart Rate as a Measure of Attentional/Cognitive Distraction

This section presents results pertaining to the seventh study objective. The 192 participants (96 in each venue) performed in-vehicle tasks while seated in a vehicle. Heart rate was measured continuously through each task using a heart rate monitor with EKG sensors attached to the torso and collarbone. The performance metric was the mean HR (number of beats per minute, BPM) for each task. HR was collected in both venues, simulator and non-driving.

3.5.1 Heart Rate: Simulator Venue

First, mean HR in the driving simulator venue was examined during the two A-V digit-recall tasks that represented unacceptable load (i.e., 2-back), acceptable load (i.e., 1-back), and no load other than the task of driving (i.e., baseline). These tasks were examined because they are benchmark tasks for different levels of cognitive load. Figure 89 shows the mean HRs during the 1-back, 2-back, and baseline trials, with one value computed for each task for each participant. Mean HRs during 2-back (*M* = 79.37, *SD* = 13.41) were higher than both 1-back (*M* = 78.4, *SD* = 13.32) and baseline (*M* = 74.54, *SD* = 12.18). T-tests showed that mean HR was sensitive to changes in cognitive load (2-back vs. baseline: *t*(206) = 2.72, *p* = .007, *d* = 0.38; 1-back vs. baseline: *t*(206) = 2.18, *p* = .03, *d* = 0.30). The difference between 1-back and 2-back mean HR was not statistically significant, *t*(206) = -0.52.

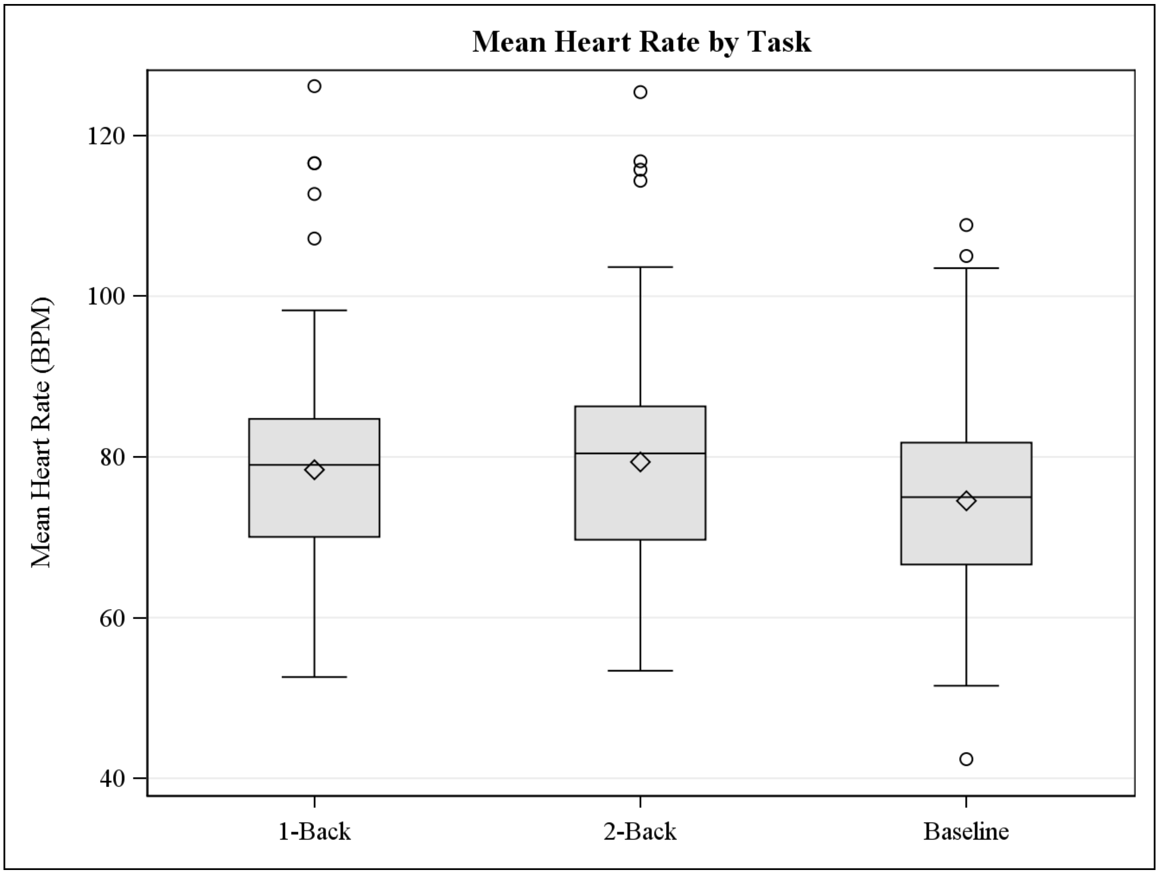


Figure 89. Mean Heart Rate by Task (Simulator venue)

Next, standard deviations of mean HR in the driving simulator was examined to attain a measure of heart rate variability (HRV). HRV is the variation between consecutive heart beats (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). The mean HR standard deviations are shown in Figure 90. HRV during 2-back (*M* = 3.43, *SD* = 1.47) significantly increased compared to the baseline condition (*M* = 2.95, *SD* = 1.29), *t*(206) = 2.53, *p* = .01, *d* = 0.35. The HRV increased during 2-back compared to 1-back (*M* = 3.09, *SD* = 1.21), but the t-test did not reach statistical significance, *t*(206) = -1.85, *p* = .07, *d* = 0.26. The t-test comparing HRV during 1-back and baseline was not statistically significant, *t*(206) = 0.81, *d* = 0.11.

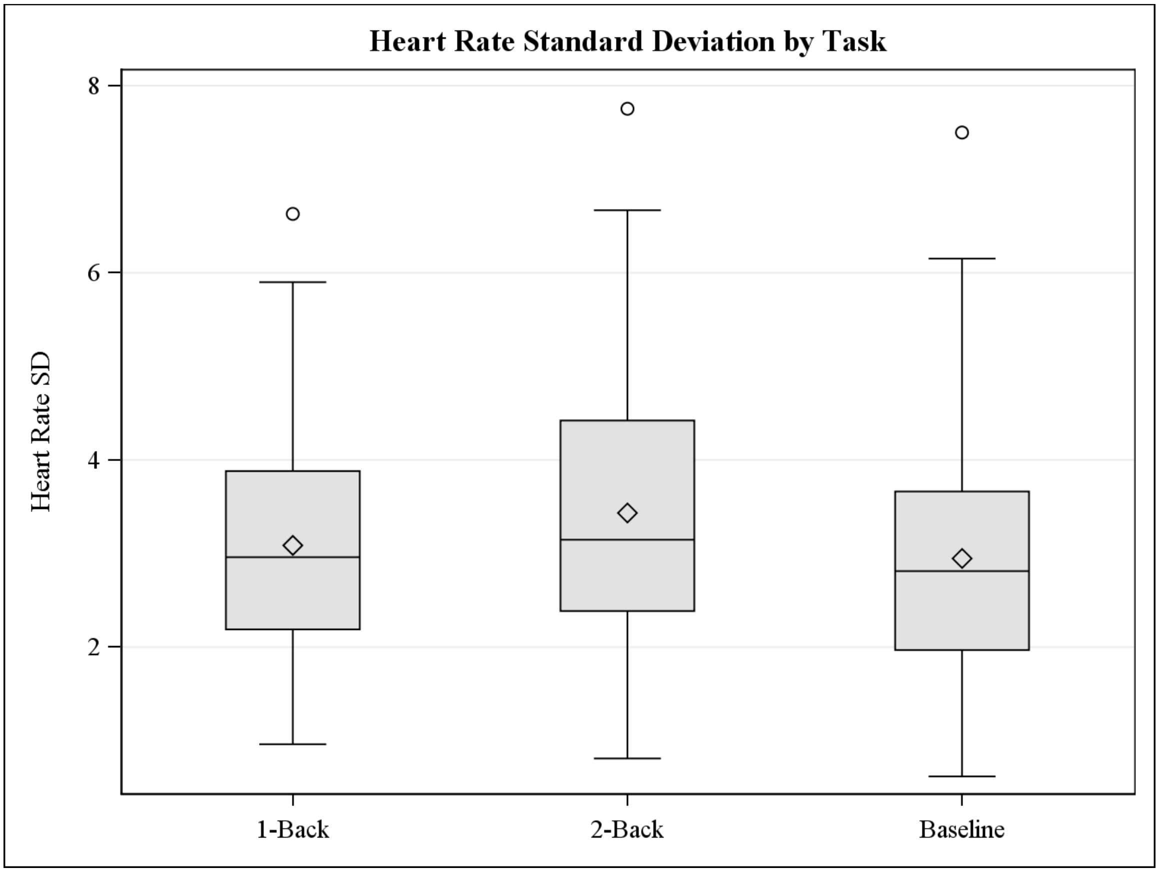


Figure 90. Heart Rate Standard Deviation by Task (Simulator venue)

3.5.2 Heart Rate: Non-Driving Venue

Figure 91 displays boxplots of mean HRs in the non-driving venue during 1-back, 2-back, and baseline tasks.

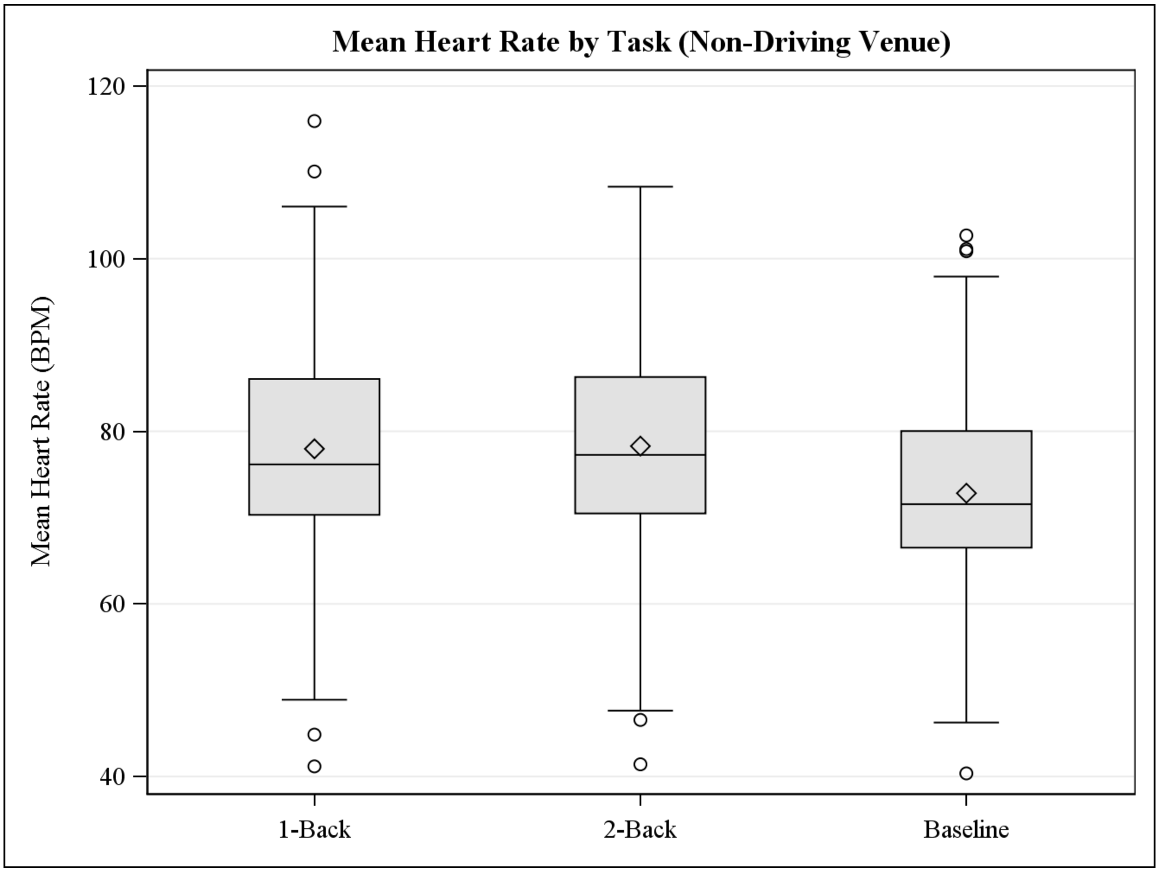


Figure 91. Mean Heart Rate by Task (Non-Driving venue)

In the non-driving venue, the no-load condition (baseline: *M* = 72.82, *SD* = 11.88) had significantly lower mean HRs compared to the medium cognitive load (1-back: *M* = 77.96, *SD* = 13.46) and high load (2-back: *M* = 78.28, *SD* = 13.2) conditions: 2-back vs. baseline, *t*(186) = 2.98, *p* = 0.003, *d* = 0.43; 1-back vs. baseline, *t*(186) = 2.78, *p* = 0.006, *d* = 0.41. HRs during the 1-back and 2-back tasks were not statistically significant, *t*(186) = -0.16, *d* = 0.02.

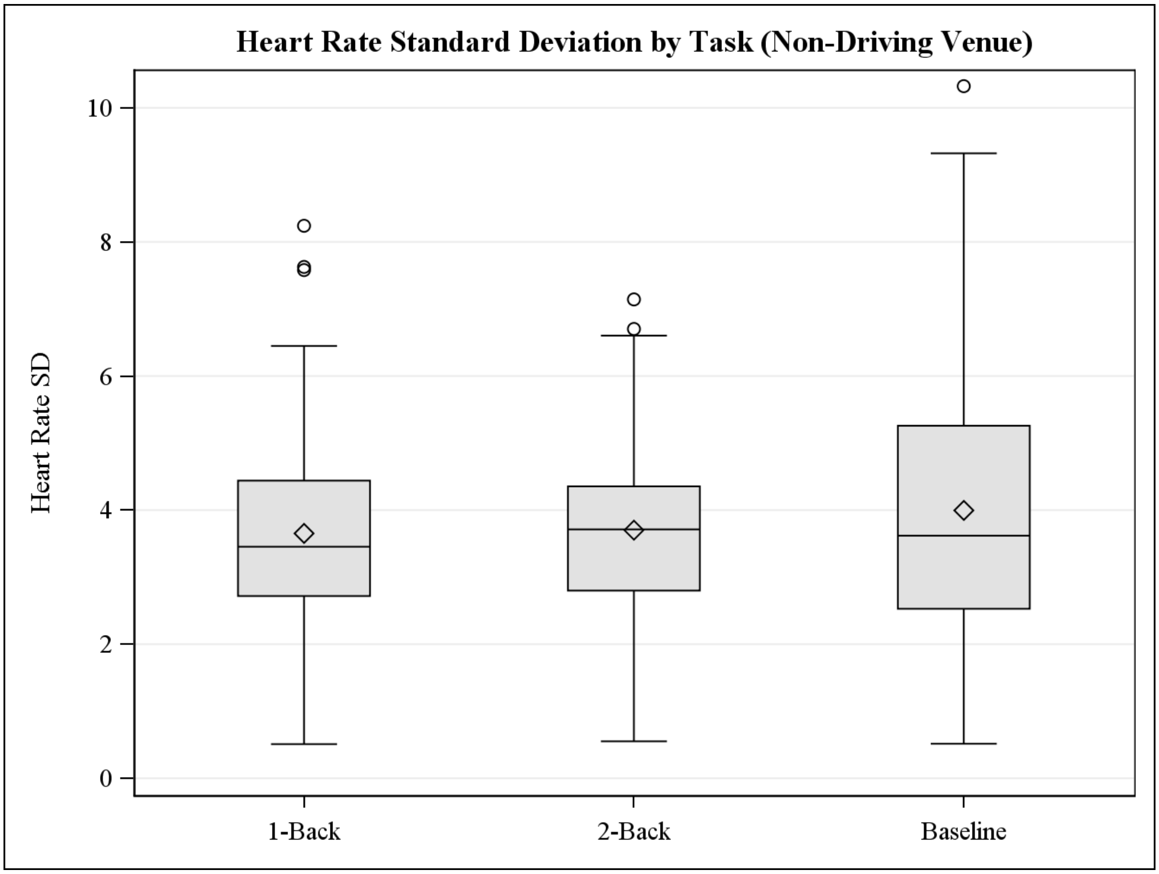


Figure 92. Heart Rate Standard Deviation by Task (Non-Driving Venue)

The boxplots for non-driving venue mean HR standard deviations are shown in Figure 92. Again, t-tests were performed to determine if there were HRV differences among tasks (i.e., 1-back, 2-back, and baseline). The 2-back (*M* = 3.70, *SD* = 1.26) and 1-back (*M* = 3.65, *SD* = 1.45) tasks showed reduced HRV compared to the baseline trials (*M* = 4, *SD* = 2). None of the t-tests for each task pairing was statistically significant (2-back vs. Baseline: *t*(186) = -1.21, *d* = 0.18; 1-back vs. Baseline: *t*(186) = -1.35, *d* = 0.20; 2-back vs. 1-back: *t*(186) = -0.25, *d* = 0.04).

3.5.3 Summary of Heart Rate Findings

1. In both venues, mean HR was sensitive to changes in cognitive demand associated with N-back (i.e., baseline, 1-back, & 2-back). This is consistent with previous findings that HR increases as demand and workload increases (Reimer & Mehler, 2011; Mehler, Reimer, & Coughlin, 2012).
2. In the simulator venue, the 2-back task significantly increased HRV compared to the no-load baseline trials. HRV during 2-back was also greater than the 1-back task, and 1-back was not significantly different than the baseline trials. In the non-driving venue, HRV was not sensitive to changes in cognitive demand.

4.0 DISCUSSION AND CONCLUSIONS

Assessing the attentional demands of auditory-vocal secondary task performance with in-vehicle devices while driving necessitates the use of additional metrics beyond those specified in the test procedures of the NHTSA Driver Distraction Guidelines. Empirical work supporting the ISO DRT standard (ISO, 2014) has established a connection between the attentional load of in-vehicle tasks and DRT performance. The present study assessed two DRT variants (TDRT and RDRT) in two test venues (i.e., driving simulator and non-driving venue). The NHTSA Driver Distraction Guidelines age and gender recommendations were used to select 192 participants. Forty-eight participants were assigned to each combination of DRT and venue. Participants performed six in-vehicle tasks, which included three discrete auditory-vocal tasks (i.e., destination entry, phone dialing, and auditory-vocal radio tuning), one discrete visual-manual task (i.e., Visual-Manual Radio Tuning [VMRT]), and two continuous auditory-vocal digit-recall tasks (i.e., 1-back and 2-back). Although there currently is no established maximum acceptable level of attentional load for tasks performed while driving, previous NHTSA-sponsored work (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014) supported the conclusion that 2-back represents a level of attentional load that is unacceptable for combination with driving and 1-back represents an acceptable level of attentional load. These tasks were included to represent benchmarks of unacceptable and acceptable levels of attentional load. The remainder of the section is organized by the specific study objectives.

Compare selected DRT variants using tasks performed with auditory-vocal interfaces

Analyses were conducted to assess differences between the test venues and DRT variants. The results identified differences between DRT variants. In the simulator, the TDRT was associated with slower responses, more variable responding across participants and more errors than the RDRT. These differences were observed on trials with in-vehicle tasks but not on baseline trials. Comparable differences between DRT variants were not observed in the non-driving venue. This pattern of results suggests that drivers in the simulator intended to respond to DRT stimuli, but one or more differences between DRT stimuli (i.e., vibrating signal on shoulder versus LED activated in the center of the visual field) led to deterioration of responses to the vibrating signal while driving and performing in-vehicle tasks. Among the possible explanations were that participants had difficulty sensing the vibrating stimulus while driving. The vibration included a buzzing sound, which may have been masked by the combination of simulated road noise, voice commands, and auditory system feedback while driving. A second possibility is that the physical requirements of driving, such as steering or reaching to touch the in-vehicle controls, reduced participants' sensitivity to the physical vibration. A third possibility is that as participants became fully engaged in driving and in-vehicle task performance, they may have inadvertently and unconsciously decided that responding to a vibrating shoulder was not essential because it has no natural association with the requirements of safe driving during in-vehicle task performance. The visual stimulus may have been more difficult to ignore, because it was presented at a location important to safe vehicle control and has a more natural association with safe driving. During the testing, some participants mentioned that they found the vibrating stimulus to be annoying. In the context of testing, slower responding alone is not a big concern. However, relatively high proportions of missed tactile signals in the simulator and the higher variability of mean RT values lead to the conclusion that the tactile DRT is not a good match for use with the simulator configuration used in this research.

Determine whether the use of the DRT provided consistent results in driving simulator and non-driving test venues

Differences between the test venues were also found. The ordering of task conditions based on mean RT values was different in the two test venues. In the non-driving venue, the three auditory-vocal tasks had DRT RT values that were considerably faster than those associated with the benchmark tasks and the VMRT. In the simulator, these three auditory-vocal tasks had mean RT values that were much closer to the benchmark-condition values. This pattern of results suggests that participants may have adopted different strategies for task performance in the different venues. In the non-driving venue, participants had more freedom to perform tasks without interruption than in the driving simulator, which required intermittent task performance due to the concurrent requirement of maintaining car-following headway. Continuous performance of in-vehicle tasks in the non-driving venue led to faster DRT RT values, indicative of lower levels of attentional demand. Intermittent task performance in the driving simulator led to slower RTs, indicative of higher levels of attentional demand. These results indicate that the relative levels of attentional load associated with a set of tasks is likely to differ by test venue. Although the present study did not include on-road driving, the results support the conclusion that the outcomes of testing performed in a non-driving venue may generalize to situations in which drivers can perform tasks continuously without disruption, but not necessarily to driving situations. DRT simulator results may be more likely to generalize to on-road driving.

Determine whether the visual metrics (i.e., occlusion and eye-glance measures) specified in the NHTSA Driver Distraction Guidelines could also be effectively used to assess auditory-vocal tasks

One objective of the study was to determine whether the NHTSA Driver Distraction Guidelines metrics can also be used to assess auditory-vocal tasks, performed with voice commands and auditory system responses, but that also require drivers to look at in-vehicle displays to obtain information or confirm that the system correctly understood voice commands. The results confirm that occlusion can be used to assess visual-manual tasks, however, they raised several questions about occlusion and voice-based tasks. One question is whether glances to in-vehicle device are necessary for successful task performance. In-vehicle systems may provide redundant feedback in response to voice commands, including auditory beeps, auditory commands, and visual displays. When participants are new to a system, as is required by the Guidelines testing protocol, they may be more likely to look to see that their commands have been accepted as they intended. In real-world use, drivers may grow to trust the systems over time and thus require fewer check glances during task performance. In simulator testing situations, particularly with low driving task demands, participants may look to the display unnecessarily because it is more interesting than the driving situation. Generally, the test protocols will only be successful if they can ensure that glances included in either TSOT (i.e., from occlusion) or TEORT (i.e., from simulator driving) values are necessary for and not incidental to task performance.

The comparison of TEORT values obtained from eye-glance data with TSOT values obtained from occlusion helps demonstrate the limitations of the occlusion protocol. Occlusion is most effective for tasks that require constant visual engagement for their performance. An assumption of 50 percent visual engagement time is built into the occlusion task structure. As shown by correlations between task-completion time and TEORT as well as by the proportions of task-completion time accounted for by TEORT, tasks that can be performed with voice commands and auditory-system responses do not satisfy this structural assumption. Thus, while driving-simulator test results may provide accurate estimates of the visual demand associated with the auditory-vocal tasks, occlusion will provide invalid results for tasks that do not require constant visual engagement.

Finally, the rigid time constraints inherent in the occlusion protocol may create problems if occlusion is used to assess tasks with potentially conflicting system-defined time constraints. The auditory-vocal phone task used in this study gives the option of visually confirming that the system correctly understood the intended phone number, but the system is relatively slow in reading back the digits that make up the dialed phone number. Thus, if a participant wanted to verify visually that the number had been correctly verbalized, it might be most efficient to wait until the system finishes its auditory presentation of the digits before looking. In driving, the participant would have the freedom to wait, but not in the occlusion protocol, which counts shutter-open time continuously during task performance. Any task that has timeout or system delays while the system processes a voice command that might cause a driver to pause and wait for the system before looking away from the roadway, will create problems for occlusion, leading to invalid estimates of visual time required for task performance.

Analyses were also conducted to assess the effects of increasing driving task demands on participants’ eye-glance behavior. In the simulator, two different car-following tasks were used, including the Guidelines-specified constant-speed car-following task, which has minimal task demand, and a higher-demand complex car-following task, which had constantly changing lead-vehicle speed together with a shorter headway that was more strictly enforced with real-time auditory feedback. TEORT distributions revealed small but consistent differences between the two car-following task conditions, both in the mean values and in the ranges of TEORT values. As predicted, TEORT values were smaller and less variable in the complex condition than in the constant-speed condition. Differences in TEORT were most pronounced for the VMRT task, which required proportionately more off-road glance behavior than the other tasks, which were performed with voice commands. Although not statistically significant, the findings indicated that increasing car-following task difficulty led to shorter and less variable TEORT values, reflecting either fewer and/or shorter off-road glances during task performance relative to the low-demand constant car-following task specified in the guidelines. The manipulation used in this study effected small changes in glance behavior particularly in changing the distribution shape, but was not strong enough to move the entire distributions of glance metrics.

Determine whether a proposed benchmark criterion level of acceptable attentional load could be established for auditory-vocal secondary tasks

One study objective was to identify a threshold to represent the upper limit of acceptable attentional demand associated with auditory-vocal secondary tasks performed while driving. One approach involved determining whether the visual-manual radio tuning task specified in the AAM Guidelines (2006), which was used to establish the guidelines visual metric criteria, could serve the same purpose with respect to attentional demand. One significant problem associated with this approach is that radio tuning tasks have changed considerably from those that were used to establish the visual-manual criteria values. There is no readily available collection of older radio tuning tasks that could be used to establish the criteria values for attentional demand. And even if such a collection were available, it is not clear how the metrics based on this collection would be used as there was no suggestion in the AAM Guidelines that their radio tuning task was intended to represent the limit of attentional demand. Rather it was intended to represent the limit of visual-manual demand. The VMRT used in this experiment is a good example of a task that is fundamentally different from the task specifications in the AAM Guidelines.

The second approach taken to identify the upper limit of acceptable attentional demand involved using the 1-back and 2-back task conditions as benchmarks. Based on previous related NHTSA-sponsored work (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014), the level of demand associated with 1-back is generally considered acceptable in relation to secondary task performance while driving, while that associated with 2-back is generally considered unacceptable. Based on analyses of group means, the accumulated body of existing research is generally consistent in showing that DRT RT is slower while performing 2-back than when performing 1-back. The present results were generally consistent with this finding, however an examination of individual performance revealed that a significant percentage of participants' DRT performance did not differ between 1-back and 2-back conditions. This led to the conclusion that creating a metric that could be used to assess individual participants' performance using the difference between 1-back and 2-back conditions is not feasible. However, the generally consistent difference between group means between 1-back and 2-back conditions reflecting the difficulty performing the 2-back task observed among most participants provides additional support for the conclusion that 2-back represents an unacceptable level of attentional demand. Accordingly, 2-back could be used as the basis for a decision model that requires acceptable tasks to have a lower level of attentional demand than 2-back. While the current study results do not permit determination of how much less attentional demand than 2-back would be acceptable, the overall difference between 1-back (which is considered to represent an acceptable level of attentional demand) and 2-back was approximately 10 percent in this study. This suggests that a target of 10-percent less than 2-back may be a candidate for such a threshold. Structural differences between auditory-vocal tasks and 2-back require further exploration as does the effect of RDRT with auditory-vocal tasks.

Assess the consistency of test results over repeated testing with multiple Guidelines groups

Observed differences in the ordering of DRT RT means for different tasks indicates that test results could differ among test groups constructed following Guidelines criteria for DRT testing. In addition, the distributions of Mean TSOT values from occlusion testing revealed differences among Guidelines groups. Specifically, the ranges and variability of TSOT values differed by Guidelines group, which indicates that test results could have been different among different Guidelines groups for tasks with mean TSOT values close to the Guidelines criterion value. These results support the conclusion that test results would be more reliable if multiple samples were used.

Establish a connection between DRT response time and brake response time (BRT) delays in emergency scenarios

Inclusion of a BRT protocol was intended to help establish a link between DRT metrics and safety. Slowed braking, particularly in emergency situations, has a more direct connection to safety than does slowed DRT responding. To the extent that BRT response can be shown to be affected by differences in attentional load among in-vehicle tasks in the same way that DRT RT is affected, it can be argued that both metrics measure the same behavioral constructs. Despite differences between DRT variants and test venues noted above, DRT RT was shown to be sensitive to differences in attentional load among tasks in the present study. DRT RT values associated with 2-back were slower than those associated with 1-back. Both tasks had slower RT values than were observed in the baseline condition. These differences reflect delays due to increased attentional load.

The BRT protocol included two types of braking events, one involving expected lead-vehicle braking and one involving the lead vehicle coming to a full stop unexpectedly without brake-light cues. Because DRT signals were expected, it was hypothesized that BRT responses to expected lead-vehicle events would provide comparable results. Expected events are thought to engage attentional resources as drivers actively evaluate the driving situation to anticipate events requiring a response. In contrast, a fully unexpected (i.e., surprise) event presumably involves no such anticipatory decision making and responses should not be affected by the level of attentional load at the time of the event.

As predicted, responses to the expected brake light activation were faster than those associated with the unexpected events. However, contrary to the study hypothesis, the RTs for expected lead-vehicle braking events were not affected by differences in attentional load. Mean BRT values in 1-back, 2-back, and baseline conditions were no different. In contrast, BRTs to the first unexpected event were affected by attentional load, inferred from differences between 2-back and baseline conditions. These results were not consistent with recent theoretical work. However, the finding that BRTs to the first unexpected event were sensitive to attentional demand in the same general way that DRT RT was sensitive to attentional demand, provides some support for the argument that DRT responses may be correlated with BRT responses in unexpected emergency situations. One caveat is necessary for interpreting this connection; because drivers controlled their own speed in the simulator, it was not possible to establish precise control of the speed and headway between vehicles at the time of the braking event. To examine the effect of headway on drivers' responses to lead-vehicle braking events, trials were separated into headway categories. Differences attributable to headway were found, but at this point, the numbers of drivers in each headway group, particularly during the first unexpected event, decreased to the point that the reliability of the mean estimates became questionable. Stronger controls of headway between vehicles at braking-event onset than those provided by the Guidelines test protocol will be necessary to test the relation between BRT and DRT RT and thus the relation of DRT performance to safety.

Participants assigned to the non-driving venue also performed in-vehicle tasks with the Occlusion protocol from the NHTSA Driver Distraction Guidelines. Participants completed five occlusion trials for each of four discrete in-vehicle tasks. Occlusion was not used with the benchmark conditions because they required continuous performance and were thus not consistent with NHTSA's testable task definition. Total shutter-open time was computed for each trial and analyzed, first in the aggregate and then within the Guidelines test protocol. When considered in the aggregate, both radio tuning tasks (i.e., visual-manual and auditory-vocal) appeared to require less than 12 seconds of TSOT when error-free trials were considered;

however, TSOT values for trials involving at least one error were greater. Of interest was the finding that 37 percent of the VMRT trials had errors. The radio tuning task in the SUV used in the study had a control button that was very difficult to use. In addition to the high proportion of errors, many participants complained about the difficulty of performing this task correctly.

Assess the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks

Heart rate was collected continuously during all trials in both the simulator and non-driving venues. The performance metric was mean heart rate (i.e., number of beats per minute). Data from the N-back (i.e., 2-back, 1-back) and baseline trials were analyzed to assess the feasibility of using heart rate as a measure of attentional demand associated with auditory-vocal tasks. Mean heart rate was sensitive to changes in cognitive demand associated with N-back in both venues. This finding is consistent with previous findings that heart rate increases as demand and workload increases (Reimer & Mehler, 2011; Mehler, Reimer, & Coughlin, 2012). Heart rate variability revealed sensitivity to increased cognitive demand but the effects were not consistent across venues.

Generally, heart rate is a more robust physiological measure of workload compared to heart rate variability and it is less impacted by normal cardiac arrhythmias (Mehler, Reimer, & Wang, 2011).

Determine whether increasing the driving simulator test scenario task demands would influence the distributions of glance metrics used in the Distraction Guidelines testing

Increasing the driving task demands in the simulator by increasing car-following task difficulty reduced the variability associated with three glance metrics, including TEORT, mean glance duration, and proportion of long glances (i.e., > 2 seconds). Increased demand was associated with small reductions in mean values for all three glance metrics, although none was statistically significant. The results suggest that a slightly greater increase in driving task demands is likely to have a stronger effect on the glance metrics which could affect Guidelines test outcomes.

4.1 Conclusions

The study results support the following conclusions:

1. Responses to visual stimuli associated with RDRT were generally faster, more consistent, and had fewer errors than responses to stimuli associated with the TDRT in the simulator. Differences between DRT stimuli in the non-driving venue were much smaller, suggesting that either DRT could effectively be used in that venue (Objective 1).
2. The driving simulator and non-driving venues did not give consistent results for some tasks. Non-driving results appear to represent the relative difficulty among tasks performed continuously without interruption. Simulator results reflect the relative difficulty among tasks performed intermittently while driving. Simulator test results are more likely to generalize to on-road driving than non-driving test results (Objective 2).
3. DRT response time is more valuable as a metric than hit hate (accuracy), which is limited by the ceiling of perfect performance, most often in the non-driving venue; however, monitoring DRT accuracy is necessary to identify speed-accuracy tradeoffs and

noticeably poor DRT performance. Participants did not trade accuracy for speed in this experiment; however, the two DRT metrics occasionally provided inconsistent results in comparisons between conditions (Objectives 1 and 2).

4. The N-back conditions provided the strongest foundation for defining a threshold of acceptable attentional demand. The conclusion that 2-back represents an unacceptable level of attentional demand supports a decision model that could require acceptable tasks to have a lower level of attentional demand than 2-back. Determining how much better requires additional research, but the difference between 1-back (which is considered to represent an acceptable level of attentional demand) and 2-back group means is approximately 10 percent in this study, which suggests that a target of 10 percent less than 2-back may be a candidate for such a threshold. Structural differences between auditory-vocal tasks and 2-back require further exploration as does the effect of RDRT with auditory-vocal tasks (Objective 4).
5. BRT results suggest that DRT RT and BRT to unexpected lead-vehicle stopping may both be sensitive to effects of attentional load. However, the inability to precisely control speed and headway at the start of the lead-vehicle braking events in the Guidelines car-following scenario precluded a strong conclusion. Stronger controls of headway between vehicles at braking-event onset than those provided by the Guidelines test protocol will be necessary to test the relation between BRT and DRT RT and thus the relation of DRT performance to safety (Objective 6).
6. Glance metrics obtained in the driving simulator (TEORT) can effectively be used to assess auditory-vocal tasks performed with voice commands and verbal system feedback or auditory-vocal tasks that also require drivers to obtain information from in-vehicle displays; however, occlusion is not suitable for use with these tasks. TSOT values obtained in the occlusion paradigm were not consistent with TEORT values obtained from glance data analysis. Occlusion is only suitable for assessing tasks with constant visual demand during task performance (Objective 3).
7. Increasing car-following difficulty led to small effects that were not statistically significant. The variability of all glance metrics was reduced with increasing task demands. The manipulation used in this study showed promise but was not strong enough to move entire distributions of glance metrics. A slightly greater increase in driving task demands is needed to affect glance metrics enough to influence Guidelines test outcomes (Objective 8).
8. The VMRT task, as performed in the Ford Explorer test vehicle used in this study, showed a high error rate in task performance. Specifically, 39 percent of participants (37/96) had three or more error trials. The NHTSA Driver Distraction Guidelines protocol calls for replacement of participants without a specified limiting provision. It is only after 24 participants have successfully completed the task that test users compute aggregate scores to determine whether the task should be determined to be “unreasonably difficult” based on the Guidelines criterion of 50 percent errors.

In summary, the driving simulator venue combined with RDRT had greater sensitivity and more consistency in detecting targeted differences than the other venue/DRT combinations. The results support the conclusion that 2-back represents an unacceptable level of attentional demand for tasks to be performed while driving. Lastly, results showed that increased driving scenario driving task demands allow for better control of off-road glance durations.

5.0 REFERENCES

Alliance of Automobile Manufacturers, Driver Focus-Telematics Working Group (2006). *Statement of principles, criteria, and verification procedures on driver interactions with advanced in-vehicle information and communication systems*. Southfield, MI: Alliance of Automobile Manufacturers.

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

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*,7-29.

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York, NY: Routledge.

Engström, J. (2010). The tactile detection task as a method for assessing drivers' cognitive load. In G. Rupp (Ed.) *Performance metrics for assessing driver distraction: The quest for improved road safety*. Warrendale Pa: SAE International, pp. 90-103.

Engström, J. (2014). Scenario criticality determines the effect of working memory load on brake response time (Unpublished report). Göteborg, Sweden: Volvo Technology/SAFER, Chalmers University.

Engström, J., Aust, M. L., & Viström, M. (2010). Effects of working memory load and repeated scenario exposure on emergency braking performance. *Human Factors, 52*, 551-559.

Harms, L., & Patten, C. (2003). Peripheral detection as a measure of driver distraction. A study of memory-based versus system-based navigation in a built-up area. *Transportation Research Part F, 6*, 23-36.

Hoffmann, E.R. & Mortimer, R.G. (1996). Scaling of relative velocity between vehicles, *Accident Analysis and Prevention, 28*, 415-421.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48*, 196-205

ISO CD 17488 (2014). Road Vehicles - Transport information and control systems - Detection Response Task (DRT) for assessing selective attention in driving. Committee Draft, ISO TC 22/SC13/WG8. Geneva: International Organization for Standardization.

Kelley, K. & Preacher, K. J. (2012). On effect size. *Psychological Methods, 17*,137-152.

Lamble, D., Laakso, M., & Summala, H. (1999). Detection thresholds in car following situations and peripheral vision: implications for positioning of visually demanding in-car displays. *Ergonomics, 42* (6), 807-815.

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

Mehler, B., Reimer, B., Coughlin, J., & Dusek, J. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record: Journal of the Transportation Research Board*, (2138), 6-12.

Mehler, B., Reimer, B., & Coughlin (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: An on-road study across three age groups. *Human Factors, 54*, no. 3. pp. 396-412.

Mehler, B., Reimer, B., & Dusek, J. A. (2013). MIT AgeLab delayed digit recall task (N-back). (MIT AgeLab White Paper Number 2013-4). Cambridge MA. Massachusetts Institute of Technology.

Mehler, B., Reimer, B., & Wang, Y. (2011, June). A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload. In *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 590-597).

Merat, N., & Jamson, A. H. (2008). The effects of stimulus modality on signal detection: Implications for assessing the safety of in-vehicle technology. *Human Factors, 50*, No. 1, 145-158.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, No. 9, 741-749.

Ranney, T. A., Baldwin, G. H. S., Smith, L. S., Mazzae, E. N., & Pierce, R. (2014, November). Detection response task (DRT) evaluation for driver distraction measurement application (Report No. DOT HS 812 077). Washington, DC: National Highway Traffic Safety Administration. Available at www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/812077.pdf

Ranney, T. A., Baldwin, G. H. S., Parmer, E., Domeyer, J., Martin, J., & Mazzae, E. N. (2011, November). Developing a test to measure distraction potential of in-vehicle information system tasks in production vehicles (Report No. DOT HS 811 463). Washington, DC: National Highway Traffic Safety Administration. Available at https://one.nhtsa.gov/DOT/NHTSA/NVS/Crash%20Avoidance/Technical%20Publications/2011/811463.pdf

Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. *Ergonomics*, *54*(10), 932-942.

Reimer, B., Mehler, B., Dobres, J., & Coughlin, J. F. (2013). The effects of a production level "voice-command" interface on driver behavior: reported workload, physiology, visual attention, and driving performance (MIT AgeLab Technical Report No. 2013-17A). Cambridge MA. Massachusetts Institute of Technology.

Strayer, D. L, Drews, F. A., & Johnson, W. A., (2003). Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9,* 23-52.

Stutts, J., Reinfurt, D., Staplin, L., and Rodgman, E. A. (2001): *The role of driver distraction in traffic crashes, Phase I Final Project Report*. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17,* 354-381.

US-EU Bilateral ITS Technical Task Force. (2010). *European Commission Expert Focus Group on Driver Distraction: Definition and Research Needs: Digital Agenda for Europe: A Europe 2020 Initiative*. Brussels, Belgium: European Commission. Retrieved from http://ec.europa.eu/digital-agenda/en/news/september-2011-driver-distraction-definition-and-research-needs-report-expert-focus-group-us%E2%80%90eu

Van der Horst, A. R., & Martens, M. H. (2010). The peripheral detection task (PDT): On-line measurement of driver cognitive workload and selection attention. In G. Rupp (Ed.) Performance metrics for assessing driver distraction: The quest for improved road safety. Warrendale, Pa: SAE International, pp. 73-89.

Victor, T. W., Engström, J., & Harbluk, J. L. (2009). Distraction assessment methods based on visual behavior and event detection. In M.A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction: theory, effects and mitigation* (pp. 135-165). Boca Raton, FL: CRC Press.

Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. 78 Fed Reg. 81, 24818, April 26, 2013. Available at www.govinfo.gov/content/pkg/FR-2013-04-26/pdf/2013-09883.pdf

Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices; Clarifications. 79 FR 55530. September 16, 2014. Available at https://www.govinfo.gov/content/pkg/FR-2014-09-16/pdf/2014-21991.pdf

Wang, J-S., Knipling, R. R., & Goodman, M. J. (1996): The role of driver inattention in crashes: New statistics from the 1995 Crashworthiness Data System. *The 40th Annual Proceedings of the Association for the Advancement of Automotive Medicine*, October 7-9, 1998, Vancouver, British Columbia.

Wierwille, W. W. (1993). *Visual and manual demands of in-car controls and displays*. Abingdon, United Kingdom: Taylor & Francis.

Wierwille, W. W., & Tijerina, L. (1998). Modeling the relationship between driver in-vehicle demands and accident occurrence. In A. Gale, et al. (Eds.), *Vision in Vehicles VI* (pp. 233-243). Amsterdam: Elsevier.

Appendix A Recruitment Material

Advertisement:

Receive $40 per hour, plus mileage allowance, for up to 4 hours of participation
We are seeking participants for a simulator study of driving performance
The study will be conducted by:
Transportation Research Center Inc. for the National Highway Traffic Safety Administration (NHTSA) of the U.S. D.O.T.
At the proving ground in East Liberty, Ohio
Weekday Sessions
MUST BE:
Licensed driver 18-70 years old
Good general health
3,000+ miles driven per year
Cell phone user while driving

PLEASE CALL: 7:00 am – 5:00 pm weekdays
[Contact name] [redacted] ext [ext]
Or [redacted][direct line #]
OR REPLY ONLINE: http://www.trcpg.com/[pagename]/
OR E-MAIL: vrtc.webmaster@dot.gov

Project Summary (to appear on TRC website):

The Transportation Research Center Inc. is conducting a research study for the United States Department of Transportation's National Highway Traffic Safety Administration (NHTSA). The study will evaluate different tools that researchers use to measure distraction caused by in-vehicle technologies and portable devices including cell phones. Participation involves one session of approximately 4 hours. Participants will drive a driving simulator and perform in-vehicle tasks like tuning a radio or dialing on a cell phone. If selected, you will be required to come to a laboratory facility located on the Transportation Research Center Proving Grounds in East Liberty, Ohio. Participants will receive $40 per hour for participating in the study. Participants will receive a monetary travel allowance for mileage to and from our facility.

Appendix B Participant Screening Questions

Phone Screen Introduction: Use only if Internet access not available to respondent

Introduction	Thanks for expressing interest in participating in our research study!
Research Study Purpose	The study is being conducted by Transportation Research Center Inc., for the United States Department of Transportation's National Highway Traffic Safety Administration (NHTSA), to evaluate different tools that researchers use to measure distraction caused by in-vehicle technologies.
Purpose of Phone Screening	During this call, I will describe the study and gather information that can be used by the principal investigator to determine if you qualify for participation. This call will take about 10 minutes. Is now a good time?
Participation Commitment	Participants will drive a driving simulator and perform in-vehicle tasks like tuning a radio or talking on a cell phone. Participation involves one session of approximately 4 hours. If selected, you will be required to come to a laboratory facility located on the Transportation Research Center Proving Grounds in East Liberty, Ohio.
Participation Compensation	If selected, you will receive $40 per hour for participating in the study. You will also receive a travel allowance for mileage to and from your home to our facility.
Information Being Requested & Confidentiality	I would now like to ask you a series of questions to determine your eligibility. Questions will cover: (1) personal information, (2) driving experience, (3) wireless phone usage, and (4) medical history. Note that we (NHTSA and TRC Inc.) will not release any personal identifying information that you provide during this call. The information gathered will be kept confidential, and stored in a password protected database on a local computer. Responses to health-related questions will be maintained separately from your personal information and will be deleted at the end of the study. You do not have to answer any question that you do not want to answer and you may end this phone call at any time. At this time, are you willing to proceed with the questions? (If YES, then proceed. If NO, then make note to delete information and inform caller that his/her information will be deleted, or if using paper survey – the survey will be shredded.)
NOTE: (Office Use Only)	Exclusion Criteria are on Subject Info Sheet.

Contact Information and Questions for Determining Initial Eligibility (Internet or Phone Interview)
Respondent Number:
Date Interviewed:
Interviewers initials:
NAME (first, last)
GENDER (M/F)
AGE (office note: must be 18 - 70 years old to participate) **18 – 70**
Have you participated in a previous driving simulator research study? If yes, what year?
How did you hear about our research study? Which newspaper or online ad?
Do you have a valid US driver's license? (Y/N)
Are there any restrictions on that license? (Y/N)
Are you able to drive without the use of assistive devices? (Y/N)
How many miles do you drive per year (office note: > **3,000**)?
What kind of vehicle do you normally drive (year, make, model)?
PHONE – daytime:
E-mail Address:
Home Address: Street Address: Apartment No:
City: State:
Zip Code:
End questions here if: (1) age outside range, (2) driver's license not valid, (3) license restrictions other than corrective eyewear, (4) miles driven per year less than 3,000 (5) previous simulator experience within last 2 years.
Secondary Questions for Determining Eligibility and Availability (Internet or Phone Interview)
Do you wear prescription eye glasses or contacts while driving? (Y/N)
Do you require reading glasses to use a cell phone while driving? (Y/N)
How comfortable (on a scale of zero to ten, with zero being least comfortable) are you at multi-tasking while driving (e.g., eating, drinking, changing radio stations, talking on a cell phone, talking with passengers)?
Do you use a cellular phone while driving? (Y/N)
Do you regularly communicate using text messages (not during driving)? (Y/N)
Do you use a navigation system, computer, or any other similar devices in your car? (Y/N)
I am now going to ask you some confidential questions about your medical history and present condition. You can refuse to answer any question. If you choose to answer, please answer YES or NO to the following.
Do you have any health problems that affect driving? (Y/N) Please describe.
Do you have high blood pressure that is not controlled by medicine? (Y/N)
Do you have a history of seizures or epilepsy? (Y/N)
Are you susceptible to motion sickness? (Y/N)

Do you have any difficulty hearing and understanding normal conversation? (Y/N)
Do you have any inner ear, dizziness, vertigo, or balance problems? (Y/N)
Do you have diabetes for which insulin is required? (Y/N)
Have you ever had a concussion, brain injury, or other injury resulting in decreased motor control or cognitive ability? (Y/N) (If yes, please describe.)
Are you taking any medications (over-the-counter or prescription) that may cause drowsiness or impact your driving ability? (Y/N)
Do you currently have any medical condition that might affect your ability to concentrate while driving, such as Attention Deficit Hyperactivity Disorder (ADHD), depression, anxiety, or claustrophobia? (Y/N) (If yes, please describe.)
Availability
Please indicate which days of the week you are available to participate in this study.
Are you available on short notice to participate in our study? Could we call you on the same day to schedule if necessary?
Can we use your email address to help with scheduling? (Y/N)
Can we use text messaging to help with scheduling? (Y/N) If yes, obtain Cell Phone number:
How long would you like to be considered for this study?
Phone Interview Wrap-up: TO END CALL SAY:
OK, that's all of our questions. Thank you! We will call you back soon and let you know if you have been accepted for participation. We will provide additional information at that time.
Phone Conversation for SCHEDULING PARTICIPANTS
The Principal Investigator or his designated associate will determine which respondents are selected for participation.
Office Use Below
WHO CONTACTED THE SUBJECT? (personnel name)
APPOINTMENT CONFIRMATION CALL BACK
Date Scheduled:
Assigned Subject Number:
Hi this is _________ from TRC Inc. This is a call back to notify you that you have been selected to participate in our driving study discussed in earlier phone conversations with ___________. I have several additional questions and then I will schedule a test session.
Since your initial screening, have you begun taking any OTC medications that might affect driving? (Y/N)

I would like to schedule an appointment with you at this time. The first available openings are: ____________ (calendar of events needed w/ date and time frame up to 4 hours - try to utilize previous call input for choices, before calling). Do any of those dates and times work for you? (If YES, schedule. If NO, then offer next available set of times, perhaps by week, until scheduled. If no good dates, find a time when best for them and say we will see what we can do and call back later.)
(Appointment Confirmation) Ok. I have you scheduled for _________________. Please try to arrive promptly.
Please be sure to bring your valid, U.S. driver's license to the appointment for identification purposes. Dress comfortably for driving and weather conditions and wear comfortable driving shoes. Do not bring another guest with you, unless prior arrangements have been made with us.
The session will last between 3 and 4 hours. You will have one or more short breaks, during which you can purchase food and soft drinks from vending machines in our lunchroom. We will have light dry snacks available at no charge.
Note that your personal wireless devices must be turned off while you are participating in this study. Cameras, firearms, and alcoholic beverages are not permitted at the data collection facility.
If your hair hangs in your face, you will be asked to pull it back out of the way so that we can see where you are looking during the experiment.
Please refrain from drinking alcohol or taking non-prescription drugs for at least the 24 hours preceding the session.
Do you understand these requirements? (Record "Yes" or "No")
DIRECTIONS: We will send a map link to you using the email address that you provided.
Do you have any questions at this time?
If you have any questions before your scheduled date, please feel free to call me at [redacted] ext. xxx. If you need to contact us on the day of your scheduled appointment, please call xxx-xxx-xxxx. Or you can send a text message to this number XXX-XXX-XXXX.
If something comes up and you need to cancel or reschedule your appointment, please try to call at least 24 hours in advance. Otherwise, we look forward to seeing you on (date at time) ___________.

Appendix C Participant Information Summary and Confidential Information Form

STUDY TITLE:	Detection Response Task (DRT) Evaluation for Driver Distraction Measurement Application: Experiment 2
STUDY INVESTIGATOR:	Thomas A. Ranney, Ph.D.
STUDY SITE:	Transportation Research Center Inc. 10820 State Route 347 East Liberty, OH 43319
TELEPHONE:	[redacted]
SPONSOR:	U.S. Department of Transportation National Highway Traffic Safety Administration (NHTSA)

You are being asked to participate in a research study. Your participation in this research is strictly voluntary, meaning that you may or may not choose to take part. To decide whether or not you want to be part of this research, the risks and possible benefits of this study are described in this form so that you can make an informed decision. This process is known as informed consent. This consent form describes the purpose, procedures, possible benefits and risks of the study. This form also explains how your information will be used and who may see it.

The study investigator or study staff will answer any questions you may have about this form or about the study. Please read this document carefully and do not hesitate to ask anything about this information. This form may contain words that you do not understand. Please ask the study investigator or study staff to explain the words or information that you do not understand. After reading the consent form, if you would like to participate, you will be asked to sign this form. You will be offered a copy of the form to take home and keep for your records.

PURPOSE

This research study is being conducted by the National Highway Traffic Safety Administration (NHTSA). The purpose of this study is to evaluate the different tools that researchers use to measure the level of distraction associated with the use of “in-vehicle technologies.” The latest in-vehicle technologies provide services such as internet access, navigation information (maps and driving directions), as well as the ability to send and receive e-mails and text messages. Many in-vehicle systems allow such tasks to be performed with voice commands and auditory responses. As new in-vehicle technologies are developed and marketed, there is a concern that these systems may interfere with driving. NHTSA is conducting this research study to determine the best way to collect data (information) on the use and impact of in-vehicle technologies while driving.

STUDY REQUIREMENTS

You are being asked to participate in this research study because:

- You are 18-70 years of age,
- You are an active driver with a valid, unrestricted U. S. driver’s license (except for restrictions concerning corrective eyeglasses and contact lenses),
- You drive at least 3,000 miles per year,
- You are in good general health, and
- You have experience using a wireless phone while driving.

NUMBER OF STUDY SITES AND STUDY PARTICIPANTS

This study will take place at one research site (noted above) and will include a minimum of 144 men and women.

STUDY PROCEDURES

Before participating in this research study, you will be asked to read this Participant Informed Consent Form in its entirety. After all of your questions have been answered, you will be asked to sign this form to show that you voluntarily consent to participate in this research study.

Your participation in this research study will consist of one session lasting approximately 3-4 hours. A member of the study staff will give you detailed instructions and will accompany you at all times during your participation in this research study. The study will be conducted in two parts. During the first part of the study, you will complete approximately 20 test trials, each lasting approximately 3-4 minutes. In each trial, you will perform a combination of an in-vehicle task and a detection response task in one of two test venues. You will perform these task combinations while you are sitting in the driver's seat of a stationary vehicle. In the second part of the study, you will complete approximately 3 trials, each lasting approximately 3-4 minutes. In each trial your will perform an in-vehicle task while driving the driving a simulator. Details are presented in the following sections.

In-Vehicle Tasks:

The in-vehicle tasks used in this study will consist of manually tuning a radio using an in-vehicle system, three tasks performed using voice commands: (1) tuning the radio, (2) dialing a phone number, and (3) entering a street address, and a verbal digit recall task that involves listening to and repeating a sequence of digits.

Detection Response Tasks:

Detection response tasks (DRT) are used to measure the amount of distraction associated with an in-vehicle task. A DRT involves a timed sequence of artificial stimuli, each requiring a button-press response. You will be assigned to one of two DRT conditions in this experiment. One DRT condition will use a simple visual stimulus (light-emitting diode or LED) that is positioned at a remote location in front of the vehicle's steering wheel. The second DRT condition will use a tactile stimulus; a small electrical vibrator will be temporarily attached to your shoulder using medical tape. In both conditions, when the stimulus is activated you will respond by pressing a button on a micro-switch that is attached to your left index finger.

Heart Rate Monitoring

Heart rate has been shown to be an indicator of the workload associated with in-vehicle tasks. During the study, a heart rate monitor will be used to record your heart rate. To obtain heart rate, two heart rate (EKG) sensors will be attached to your torso; one will be placed just below the collarbone on the right side of your body; another will be placed on the opposite side of your body over the lower rib. An alcohol pad will be used to clean the two spots before attaching the sensors. This helps us get a better reading from the EKG sensors. The sensors are attached with foam tape and a sticky gel. They will be placed on the areas where the skin is cleaned. Thin wires connected to the sensors will run under your clothing to a lightweight plastic harness that will be attached to your shoulder. The harness connects the sensors to the heart rate monitor.

Test Venues:

The experiment will have two parts. In the first part, all participants will be assigned to one of two test venues, which include a "Driving Simulator" or a "Non-Driving" venue. Both venues will involve sitting in a vehicle and performing the in-vehicle tasks and the DRT. In the driving simulator test venue, these tasks will be performed together with a simulated driving task. In the non-driving venue, the in-vehicle tasks and the DRT will be performed alone, with no driving task.

In the second part of the experiment, all participants will drive the simulator in one of three in-vehicle task conditions. The purpose of this part of the experiment is to obtain driving simulator data in these three conditions without the DRT. Therefore, there will be no DRT in this part of the experiment.

Driving Simulator

The driving simulator used in this study is a fixed-base simulator. A fixed-based simulator is a system that imitates the conditions of driving in real life, but does not move. The simulator will be connected to a 2011 Ford Explorer, called the "Study Vehicle." The study vehicle will have its engine turned off. While driving the simulator, you will sit in the driver's seat of the study vehicle and control the simulator by moving the steering wheel and the gas and brake pedals of the study vehicle. The study vehicle is equipped with sensors to collect information on your steering, braking and gas pedal usage. The sensors are located so that they will not affect your driving. The information collected by these sensors is recorded so that it can be analyzed at a later time. A large screen in front of the study vehicle will display a computer-generated image of the virtual road on which you will be driving.

While operating the simulator, you will be asked to perform specific driving tasks. These tasks will involve activities such as following a car at a specified distance and keeping the vehicle within the specified travel lane.

Summary of Study Procedures:

The following procedures will take place at your session:

- After signing this consent form, you will be assigned to one of two test venues. If you are assigned to the driving simulator venue, you will be provided instructions and training on driving the simulator, DRT, and performing the in-vehicle tasks. If you are assigned to the non-driving test venue, you will be provided instructions and training on the DRT and performing the in-vehicle tasks. You will also be given the opportunity to practice each of these before performing test trials.
- You will then complete 1 set of approximately 20 trials each lasting approximately 3.5 minutes.
- After completing the first set of trials, all participants will complete one set of three approximately 3.5-minute trials in the driving simulator. You will be provided instructions and training on performing the in-vehicle tasks. In addition, if you were initially assigned to the non-driving test venue, you will be provided instructions and training on the driving simulator.
- After completing both sets of trials, the session will end and your participation will be complete.

NEW INFORMATION

No changes to procedures during this study are anticipated. However, any new information developed during the course of the research that may affect your willingness to participate will be provided to you.

RISKS of STUDY PARTICIPATION

Most people enjoy driving in the simulator and do not experience any discomfort. However, a small number of participants experience symptoms of discomfort associated with simulator disorientation. Previous studies with similar driving intensities and simulator setups have produced mild to moderate disorientation effects such as slight uneasiness, warmth, or eyestrain for a small number of participants. These effects typically last for only a short time, usually 10-15 minutes, after leaving the simulator. If you ask to stop driving as a result of discomfort, you will be allowed to stop at once. You will be asked to sit and rest. You will also be given the opportunity to consume a beverage and/or a snack. After resting, you will be given the opportunity to decide whether to continue your participation in the experiment or to leave. There is no evidence that driving ability is hampered in any way by simulator disorientation; therefore, if you decide to leave and show minimal or no signs of discomfort, you should be able to drive home. If you experience anything other than slight effects, transportation will be arranged through other means. This outcome is considered unlikely since studies in similar devices have shown only mild effects in recent investigations and evidence shows that symptoms decrease rapidly after simulator exposure is complete.

If you are assigned to the tactile DRT condition, a small electrical vibrator will be attached to your shoulder with medical adhesive tape. The vibrator will be activated periodically for durations of 1 second. The level of vibration will be set to be noticeable but not uncomfortable. The associated risk is expected to be no more than mild discomfort for a small percentage of participants.

The heart rate monitoring equipment will require attaching two sensors to your skin. Some participants may experience minor discomfort during removal and/or minor skin irritation following removal of the sensors. The associated risk is expected to be no more than mild discomfort for a small percentage of participants.

There are no known physical or psychological risks associated with participation in this study beyond these.

BENEFITS of STUDY PARTICIPATION

This research study will provide data on driver behavior and in-vehicle task performance that will be used by researchers to provide a scientific basis for developing recommendations or standards for performing in-vehicle tasks while driving.

You are not expected to receive direct benefit from your participation in this research study.

ALTERNATIVES

This study is for research purposes only. Your alternative is to not participate.

CONDITIONS OF PARTICPATION, WITHDRAWAL, AND TERMINATION

Participation in this research is voluntary. By agreeing to participate, you agree to operate the research equipment in accordance with all instructions provided by the study staff. If you fail to follow instructions, or if you behave in a dangerous manner, you may be terminated from the study. You may withdraw your consent and discontinue participation in the study at any time without penalty.

COSTS TO YOU

Other that the time you contribute, there will be no costs to you.

COMPENSATION

You will receive $40.00 per hour for the time you spend at the data collection facility. You will receive mileage reimbursement for travel to and from the data collection site.

If you voluntarily withdraw or are terminated from this study, you will be compensated for the number of hours that you participated in the study.

INFORMATION COLLECTED

In the course of this study, NHTSA will collect the following data to assess your eligibility for study participation and to document your participation:

1. **Contact information** includes your name, address, e-mail address, phone number(s), and similar information used to contact you when needed in relation to your study participation.
2. **Driving background and experience information** includes your driver's license information, number of years of driving experience, and your personal's vehicle make, model, and model year. This information is used to verify your identity and characterize your level of driving experience.
3. **Health data** includes your responses to questions about certain health conditions you may have and medications you may take that may affect your ability to drive normally.
4. **Engineering data includes** driving performance, behavior, and physiological data collected from the simulator or other systems you interact with during your participation, subjective rating data, and sensor data. This includes measures of task performance including response time, task time, number of errors made, and similar information. Sensor data includes vehicle motion information such as speed and the timing and magnitude of vehicle control inputs made including steering, gas pedal, and brake pedal inputs. Physiological data includes heart rate data.
5. **Video/audio data** (i.e., the information recorded by video cameras) including images of your face will be recorded to permit analysis of the location and duration of your eye glances while driving and performing other study-related tasks. Recorded video data containing images of your face could be used to personally identify you. Video of your face and head will include some added space around the head to compensate for any head movements. Video cameras will also capture views of your torso the forward simulated road scene or other task related images to permit analysis and characterization of your performance of tasks. Audio data will include your vocal responses to tasks performed as part of the experiment. All video/audio will be captured and stored in digital format (no tape copies will exist).

> **Any data collected during this study that personally identifies you or that could be used to personally identify you will be treated with confidentiality**. Study data (which includes Engineering and Video/Audio data as defined above) is collected without reference to your name or contact information. Study data is completely separate from your name and other identifying information and only identified with a subject number. Your name also will be separated from any data about you, either provided by you in response to questionnaires or gathered by researchers during the study, and will be replaced by the same subject number. The list correlating subject numbers to study data will be kept in a secure location and only specific people (the study Investigators) will have access for specific study-related reasons, such as to contact you for an additional participation appointment if some portion of your study data is found to have problems that require certain test trials to be repeated. Contact information data as well as the list correlating subject numbers to study data will be stored on password-protected directories and destroyed after the study is complete, unless you have indicated that you have interest in participating in future NHTSA studies in which case we will retain your name, contact information and the data provided by you in connection with screening process (except specific health information) by which we selected you to participate in this study.
>
> Study data will be securely transferred from the driving simulator or data acquisition system to secure password-protected directories and verified. Verified valid data from all study participants will be combined for analysis.

Persons who will have access to the study data that could personally identify you, including facial video, will include only authorized NHTSA personnel and other researchers authorized by NHTSA. Access to study data will be solely for authorized research purposes. Such data will be maintained only on secure computers and/or file directories that are password-protected. In consenting to this study, you also are consenting to follow-on research involving the data collected. However, any future use of your identifying data for research purposes will require approval by an Institutional Review Board. In addition, future uses of your data by researchers outside NHTSA will require data sharing agreements that provide an equal or greater level of confidentiality for your data as is provided by this agreement for this project.

> The study data collected will be stored securely in electronic form indefinitely. The study data will be kept secure through storage in a specific password-protected project folder on a secure server drive. At the conclusion of the data collection phase of this study, both raw and validated analysis datasets, without any link to your contact information, will be securely stored for a period of time specified per NHTSA's data retention policy.

USES AND RELEASE OF INFORMATION COLLECTED

Information NHTSA may release:

The engineering data collected and recorded in this study will include performance scores based on the data. This data will be analyzed along with data gathered from other participants. NHTSA may publicly release this data, which will not be linked to your name or contact information, in final reports or other publication or media for scientific, educational, research or outreach purposes.

The video/audio data recorded in this study includes your images of your face and in-vehicle audio (including your voice). The video/audio data may include information regarding your driving performance. Video and in-vehicle audio will be used to examine your driving performance and other task performance while driving. NHTSA may publicly release video image data (in continuous video or still formats) and associated audio data, either separately or in association with the appropriate engineering data for scientific, educational, research, or outreach purposes. In doing so, the video will not be linked to your name or your contact information and images of your face will be obscured.

NHTSA may show specific clips of video at research conferences. NHTSA also may show specific clips of video to the media, driver's education teachers and students, and others involved in efforts, to improve highway and road safety. In doing so, the video may include unobscured images of your face, but will not be linked to your name or contact information.

Information NHTSA may not release:

Any release of engineering data or video/audio data shall not include release of your name. However, in the event of a court action, NHTSA may not be able to prevent release of your name or other personal identifying information. NHTSA will not release any information collected regarding your health and driving record.

Other NHTSA use of information:

It is expected that the data captured throughout the course of the entire study, including that from all of the study participants, will be a valuable source of data on how drivers perform in and respond to certain situations and how vehicles might be enhanced to improve safety. Therefore, it is possible that there may be follow-on data analyses using all or part of the data for years into the future. In consenting to this study, you are consenting to future research uses of the information and video/audio data gathered from you.

QUESTIONS

Any questions you have about the study can be answered by Thomas Ranney, Ph.D., or the study staff by calling 1-800-262-8309.

If you have any questions regarding your rights as a research participant, or if you have questions, concerns, complaints about the research, would like information, or would like to offer input, you may contact: Steven L. Saltzman, M.D., Chairman of Sterling Institutional Review Board, 6300 Powers Ferry Road, Suite 600-351, Atlanta, Georgia 30339 (mailing address) at telephone number [redacted] (toll free).

INFORMED CONSENT

By signing the informed consent statement contained in this document, you agree that your participation is voluntary and that the terms of this agreement have been explained to you. Also, by signing the informed consent statement, you agree to operate the study equipment in accordance with all instructions provided by the study staff. You may withdraw your consent and discontinue participation in the study at any time without penalty.

NHTSA will retain a signed copy of this Informed Consent form. A copy of this form will also be offered to you.

INFORMED CONSENT STATEMENT

I certify that:

- I have a valid, U. S. driver's license.
- All personal and vehicle information, as well as information regarding my normal daily driving habits provided by me to NHTSA, and/or Transportation Research Center Inc. employees associated with this study during the pre-participation screening and the introductory briefing was true and accurate to the best of my knowledge.
- I have been informed about the study in which I am about to participate.
- I have been told how much time and compensation are involved.
- I have been told that the purpose of this study is to evaluate the tools that researchers use to measure driving and in-vehicle task performance.
- I agree to operate the research equipment in accordance with all instructions provided to me by the study staff.

I have been told that:

- Part of the study will be conducted in a fixed-base driving simulator and that the risk of discomfort associated with simulator disorientation is minimal.
- For scientific, educational, research, or outreach purposes, video images of my driving, which will contain views of my face and accompanying audio data, may be used or disclosed by NHTSA, but my name and any health data or driving record information will not be used or disclosed by NHTSA.
- My participation is voluntary and I may refuse to participate or withdraw my consent and stop taking part at any time without penalty or loss of benefits to which I may be entitled.
- I have the right to ask questions at any time and that I may contact the study investigator, Thomas Ranney, Ph.D., or the study staff at [redacted] or [redacted] for information about the study and my rights.

I have been given adequate time to read this informed consent form. I hereby consent to take part in this research study.

I, ____, voluntarily consent to participate.
(*Printed Name of Participant*)

____
Signature of Participant Date

INFORMATION DISCLOSURE

By signing the information disclosure statement contained in this document, you agree that the National Highway Traffic Safety Administration (NHTSA) and its authorized contractors and agents will have the right to use the NHTSA engineering data and the NHTSA video and audio data for scientific, educational, research, or outreach purposes, including dissemination or publication of your likeness in video or still photo format, but that neither NHTSA nor its authorized contractors or agents shall release your name; and you have been told that, in the event of court action, NHTSA may not be able to prevent release of your name or other personal identifying information. NHTSA will not release any information collected regarding your health and driving record, either by questionnaire or medical examination. Your permission to disclose this information will not expire on a specific date.

Information Disclosure Statement

I, ____, grant permission to the National Highway
(*Printed Name of Participant*)

Traffic Safety Administration (NHTSA) to use, publish, or otherwise disseminate NHTSA engineering data and NHTSA video image data, as defined in the Participant Informed Consent Form (including continuous video and still photo formats derived from the video recording), and associated with the appropriate engineering data for scientific, educational, research or outreach purposes. I have been told that such use may involve widespread distribution to the public and may involve dissemination of my likeness in video or still photo formats, but will not result in release of my name or other identifying personal information by NHTSA or its authorized contractors or agents. I have been told that my permission to disclose this information will not expire on a specific date.

____
Signature of Participant Date

Appendix D Participant Instructions

STUDY DESCRIPTION OVERVIEW

Thanks for agreeing to participate. Please feel free to ask questions at any time and let us know if at any point you feel you need a break.

The experiment will involve approximately 20 trials, including practice and testing. Each trial will last about 4 minutes.

In each trial, you will perform an in-vehicle task and a detection response task in one of two test venues. We will give you specific instructions before each trial. Please make sure that you don't start a trial if you are confused or don't know what we are asking you to do.

In-Vehicle Tasks:

The in-vehicle tasks, which will also be called 'secondary tasks' include: (1) manually tuning an in-vehicle radio, (2) verbally tuning an in-vehicle radio, (3) verbally dialing a phone number, (4) verbally entering an address into a navigation system, and (5) performing a verbal task that is similar to a hands-free cell phone conversation.

Detection Response Tasks:

The detection response tasks (or DRTs) involve a timed sequence of stimulus signals, each requiring a button-press response. All participants will be assigned to one of two DRT conditions. In one condition, the signal will be a light emitting diode (LED) that is positioned in front of the vehicle's steering wheel. In the other condition, the signal is a small electrical vibrator that will be temporarily attached to your shoulder (clavicle bone) using medical tape. When you see the LED illuminate or feel the vibration you should respond as quickly as possible by pressing a button that will be attached to your left index finger.

Test Venues:

The experiment will have two parts. In the first part, all participants will be assigned to one of two test venues, which include a driving simulator or a non-driving venue. Both venues will involve sitting in a vehicle and performing the in-vehicle tasks and detection response tasks. In the "Driving Simulator" test venue, these tasks will be performed together with a simulated driving task. In the non-driving venue, the in-vehicle and detection response tasks will be performed alone, with no driving task.

In the second part of the experiment, all participants will drive the simulator while performing secondary tasks. There will be no detection response tasks in this part of the experiment.

Driving Simulator:

The driving simulator used in this study is a fixed-base simulator. It does not move. The simulator is connected to a 2011 Ford Explorer. While driving the simulator, you will sit in the driver's seat. You will control the simulator by moving the steering wheel and pressing the accelerator and brake pedals of the study vehicle. The vehicle will have its engine turned off. A large screen in front of the vehicle will display a computer-generated image of the road on which you will be driving. Do you have any questions so far?

Test Venue Orientation

SIMULATOR ORIENTATION

This vehicle is a Ford Explorer, which has been modified to collect driving performance data. You will be sitting in this vehicle to drive the simulator. Please get into the driver's seat and adjust the seat to your comfort. Make sure that you can reach the buttons on the center console, for we will be using the center console to perform some secondary tasks while seated in this vehicle. The seat controls are on the lower left side of the seat. There is no need to adjust the mirrors as you will not be using them for this experiment. No shifting is required in this vehicle.

We have added sensors to the steering wheel, accelerator and brake pedals. These sensors allow us to run the driving simulator without having the vehicle turned on. Your control inputs are recorded by these sensors and used by the simulator to change the roadway image projected on the screen in front of you.

While driving in the simulator, remember, safe driving is the highest priority! You should do your best to keep your vehicle centered in the designated travel lane at all times and to maintain a constant following distance behind the lead vehicle. Car following and lane keeping performance are both measured as part of the primary task of driving. The car-following task will be explained to you in a few minutes.

Do you have the seat adjusted the way you like it?

Test Venue Orientation

NON-DRIVING VEHICLE ORIENTATION

This vehicle is a Ford Explorer, which has been modified to collect test performance data. Please get into the driver's seat and adjust the seat to your comfort. The seat controls are on the lower left side of the seat. Make sure you can reach the buttons on the center console, for we will be using the center console to perform some secondary tasks while seated in this vehicle. In this vehicle, you will not need the mirrors or vehicle controls.

Do you have the seat adjusted the way you like it?

TASK PERFORMANCE FEEDBACK DESCRIPTION

This table defines three levels of performance for car following and lane keeping in the driving simulator test venue and for the detection response and in-vehicle or secondary tasks that are used in both test venues. In the non-driving test venue, only the detection response and secondary tasks are performed. All of these tasks will be explained later, but for now we will review the general performance criteria.

Task	**Good Performance**	**Acceptable Performance**	**Poor Performance**
Car Following	Maintains following distance consistently with minor deviations	Maintains following distance mostly with some noticeable deviations	Generally fails to maintain following distance
Lane Keeping	Maintains lane position consistently with minor deviations	Maintains lane position mostly with some noticeable deviations	Generally fails to maintain lane position
Detection Response Task (DRT)	Consistently attentive to DRT detection, detecting most stimuli	Moderate number of DRT stimuli not detected	Fails to detect significant number of DRT stimuli
Secondary Tasks	Performs secondary task continuously with minimal errors	Performs secondary task either intermittently or with moderate number of errors	Performs secondary task with considerable difficulty, slowly, and with moderate number of errors

Do you have any questions about the performance feedback?

SIMULATOR DRIVING TASK INSTRUCTIONS

Our simulator is a fixed-base driving simulator, meaning that it has no motion. The simulated driving environment will be a 4-lane roadway with a lead vehicle traveling in front of you.

When the roadway image first appears, your vehicle will be stopped and you should accelerate to 50 mph. After several seconds, a vehicle will appear ahead of you in your travel lane. We call this the "lead vehicle" because it is leading you in the car following task. Your task is to follow this vehicle, adjusting your speed as necessary to maintain a constant following distance. When the lead vehicle first appears, it will be 220 feet ahead of you. This is the desired following distance. You should take note of this distance when the vehicle first appears on the screen and try to maintain this following distance throughout the entire drive. The lead vehicle will maintain a constant speed of 50 mph throughout the drive.

Remember, safe driving is your highest priority! Both car following and lane keeping performance are measured as part of the primary task of driving. You should keep the vehicle in the center of the right lane and do your best to maintain the initial following distance behind the lead vehicle. If your following distance increases, it is OK to drive faster than 50 mph to catch up to the lead vehicle. If your following distance decreases, it is OK to drive slower than 50 mph to return to the specified following distance.

On each trial, you will drive approximately 3 miles. You should continue driving and performing the secondary task until the lead vehicle disappears, which signifies the end of the trial. Shortly thereafter, the simulator screen will shut off and go blank.

Do you have any questions or need a repeat of any instructions about the driving simulator or car following tasks before we practice?

SIMULATOR AND DRIVING TASK FAMILIARIZATION DRIVE
INFORMATION FOR EXPERIMENTERS

SIMULATOR FAMILIARIZATION DRIVE (lead vehicle is not present, no CF (Fam.evt))

When participant is ready to drive simulator:

> *"This drive is your 'test drive.' We want you to get a feel for driving in the simulator. The road will be straight except for one initial curve. There will be no other traffic or in-vehicle tasks. Remember to keep your hands and feet off the controls until the roadway image appears.*
>
> *When the roadway image appears, you may begin to press the accelerator and steer the vehicle. Speed up to about 50 mph and then slow down using the brake. Try making a lane change, then try keeping the vehicle centered in the travel lane for a while. Try maintaining a constant speed. Do whatever you need to become comfortable driving the simulator."*

When drive is over,

> *"Ok. Do you have any questions or do you want to practice this drive again?"*

DRIVING TASK FAMILIARIZATION DRIVE (car following, lead vehicle (FamCF.evt))

> *"In the next drive, we will add the car following task."*
>
> *"This drive will begin like the last one, but shortly after you get around the initial curve a lead vehicle will appear ahead of you in your travel lane. Make sure you are driving at approximately 50 mph when the lead vehicle appears because that is the initial speed of the lead vehicle. Remember to make note of the distance between your vehicle and the lead vehicle when it first appears, as this is the desired following distance that you should try to maintain throughout the drive."*
>
> *"The speed limit sign says 50 mph, but you can drive faster to catch up to the lead vehicle if you fall behind. In our scoring, your ability to maintain the designated following distance is our primary performance measure. You should also drive in the right lane and try to keep the vehicle centered in that lane at all times."*

After drive, provide performance feedback:
Following distance and lane keeping performance
Repeat driving task instructions as needed

Subjects can repeat this practice drive as needed, and should repeat if they have any difficulty, such as poor car following performance:

> *"Ok. Do you have any questions about the car following and lane keeping task or do you want to practice this drive again before we move on?"*

DETECTION RESPONSE TASK (DRT) TRAINING

While performing each secondary task, you will be asked to respond to a detection response task (DRT), which requires you to respond to a sequence of simple stimuli that will be presented to you one at a time. You will respond to a stimulus by pressing a micro-switch that will be attached to your left index finger. The micro-switch is attached by wire to our data acquisition system. This equipment allows us to record the time at which each response is made.

The two detection response task types are: a tactile DRT and a remote DRT. I will show you which one you will be using and let you practice the task momentarily, but first I will explain the stimulus and response method.

[Experimenter: Read the appropriate one of the following paragraphs.]

> **Tactile DRT** – The stimulus for the tactile DRT will be a localized vibration. A small vibrator, called a tactor, will be taped to your left shoulder near the clavicle. When you feel the vibration, you should respond as quickly as possible by pressing the micro-switch attached to your finger. The vibration will occur every 3 to 5 seconds during each trial and will remain active until you press the button or for 1 second, whichever comes first. You will be scored based on your speed and accuracy in detecting the stimuli.
>
> **Remote DRT** – The stimuli for the Remote DRT will be a single red light-emitting diode (LED). When you see the LED illuminated, you should respond as quickly as possible by pressing the micro-switch attached to your finger. The LED will be activated every 3 to 5 seconds during each trial and will remain active until you press the button or for 1 second, whichever comes first. You will be scored based on your speed and accuracy in detecting the stimuli.

Now, I will show you the DRT and allow you to try it. First, please place the response button on your left index finger such that the button is comfortable and can be pressed while you are holding the steering wheel.

> [In the **Simulator** venue] Please respond to a stimulus by pressing the button against the steering wheel, and use this method (pressing against the steering wheel) consistently throughout this test venue.
>
> [In the **Non-driving** venue] You may press the button against either the steering wheel or against your thumb. Please choose one method (thumb or steering wheel) during this training and then use it consistently throughout this test venue.

[Experimenter: Make sure button and wire are positioned correctly.]

Ok, here's the first stimulus.

> Go ahead and try a few button presses in response to the stimuli. If you press the button quickly, a stimulus will shut off. If you do not respond quickly, it shuts off after 1 second.

Do you have any questions about the detection response task? You will be given the opportunity to practice the DRT again, in combination with the secondary tasks before the main trials.

SECONDARY TASK INSTRUCTIONS OVERVIEW

The in-vehicle tasks are called secondary tasks because they have a lower priority than driving. Driving is the primary task. In the simulator, car following and lane keeping are the main components of the primary task. In the non-driving test venue, there is no primary driving task, but the in-vehicle tasks are still called secondary tasks.

Instructions for performing secondary tasks will be presented auditorily so you don't have to look for this information. For some tasks, information will also be displayed on the computer screen located to the right of the center console. We call this the Task Screen.

> **[All participants**] In the simulator, the first secondary task will be presented shortly after the lead vehicle appears.
> **[Non-driving participants only**] In the non-driving test venue, the experimenter will initiate the secondary task.

You will perform secondary tasks repeatedly over trials that will last several minutes.

For some tasks, you will work continuously with only minor pauses. For these tasks, information is presented auditorily throughout the trial and you will respond verbally. No information will be presented on the Task Screen.

Other secondary tasks, like dialing a phone or tuning the radio, have well-defined beginning and end points. For these tasks, you will be given new instructions (like a different phone number or a different radio frequency) at regular intervals throughout a trial. This means that you will only have a limited amount of time to complete each task instance. Instructions will be presented auditorily. The information will also be presented on the Task Screen and will be available while you are performing the task. The pacing is designed so that you have a few seconds between task instances. However, to ensure that you finish the secondary task within the allotted interval, you cannot wait too long before starting. You also cannot spend too much time fixing errors.

Don't worry if you make an error. We don't expect perfect performance. If you make an error while performing a secondary task, please try to correct it before moving on. We will provide specific information about how to correct errors. It is important that you try to complete each task if possible.

Do you have any questions or need a repeat of any instructions before we move on to training for today's first secondary task?

VISUAL-MANUAL RADIO TUNING INSTRUCTIONS (1)

In this task you will tune the radio to a designated frequency by using the touch screen and the tuning controls located below the center console. During the trial, you will select several different radio frequencies, one at a time. The bands (AM or FM) and frequencies will be presented auditorily and will be visible on the Task Screen.

> **[All participants**] In the simulator venue, the first frequency will be presented shortly after the lead vehicle appears.
> **[Non-driving participants only**] In the non-driving venue, we will begin the trial when you are ready.

At the beginning of a trial, you will hear the first frequency followed by the word “Begin.” At this point, you should work quickly and accurately to complete the task. If you forget the frequency, you can look at the Task Screen. The information will be presented there.

[Experimenter: Make sure we start on radio/audio screen.]

First, select the frequency band by pressing the AM or FM button located to the left on the touch screen. The current band will be highlighted in red with the current frequency in large font on the screen. If you select the wrong band, press the correct button for the appropriate band.

Use the tuning controls located below the screen (**Tune +** and **Tune -**) to adjust the frequency. When you have reached the specified frequency, say ‘DONE’ aloud to indicate that you have completed the radio tuning task.

If you select the wrong band or frequency, try to fix it before moving on. If you notice an error but have already said ‘DONE’, you do not need to try to fix it. The task is considered complete when you say ‘DONE’.

Once the task is complete, you will wait for the next frequency to be presented. After you hear the new frequency and the word ‘Begin’, you will then perform the same sequence, starting with the band selection. You will continue in this way until the trial is complete.

Do you have any questions or need a repeat of any instructions before we practice this task?

[Load RadioTrain.tsv for stationary practice.]

AUDITORY-VOCAL RADIO TUNING INSTRUCTIONS (2)

In this task you will tune the radio to a designated frequency by using voice commands. During the trial, you will select several different radio frequencies, one at a time. The bands (AM or FM) and frequencies will be presented auditorily and will be visible on the Task Screen.

> **[All participants**] In the simulator venue, the first frequency will be presented shortly after the lead vehicle appears.
> **[Non-driving participants only**] In the non-driving venue, we will begin the trial when you are ready.

At the beginning of a trial, you will hear the first frequency followed by the word “Begin.” At this point, you should work quickly and accurately to complete the task. If you forget the frequency, you can look at the Task Screen. The information will be presented there.

First, press the “Push to Talk” button on the right side of the steering wheel, and listen for the system prompt to indicate it is listening for your voice command.

When the system is listening, you should state the band and frequency you were given. This information is available on the Task Screen if you need it. After saying the band and frequency aloud, you should listen for the system prompt that will confirm your selection. If the system has verbally confirmed the correct frequency, please say “DONE” aloud to indicate that you have completed the task.

If the system selected the wrong frequency, you should repeat the process by pressing the “Push to Talk” button and repeating the band and frequency aloud, which will still be visible on the task screen. If the system tunes to an incorrect frequency a second time, you should go ahead and just say ‘DONE’ aloud to make sure that you have a few seconds to prepare for the next frequency to be presented. In other words, after two attempts the task will be considered complete.

Once the task is complete, you will wait for the next frequency to be presented. After you hear the new frequency and the word “Begin,” you will perform the same sequence, starting with the “Push to Talk” button. You will continue in this way until the trial is complete.

Do you have any questions or need a repeat of any instructions before we practice this task?

[Load RadioTrain.tsv for stationary practice.]

AUDITORY-VOCAL PHONE DIALING INSTRUCTIONS (3)

In this task you will dial a phone number by using voice commands. During the trial, you will dial several different phone numbers, one at a time. The phone numbers will be presented auditorily and will be visible on the Task Screen.

> **[All Participants**] In the driving simulator, the first phone number will be presented shortly after the lead vehicle appears.
> **[Non-Driving Participants Only**] In the non-driving venue, we will begin the trial when you are ready.

At the beginning of a trial, you will hear the first phone number followed by the word “Begin.” At this point, you should work quickly and accurately to complete the task using voice commands.

First, press the “Push to Talk” button on the right-hand side of the steering wheel, and listen for the system prompt to indicate it is listening for your voice command.

When the system is listening, you should say ‘Dial’, and then wait for the system to prompt you to say the phone number. Then, say the phone number aloud one digit at a time. Use a normal pace without pausing. Saying all the numbers together is important. If you pause, the system may get confused and not understand the numbers you say after you pause.

When you have completed entering the phone number, the system will repeat the digits that you have entered. They will also appear on the system display. If the phone number is correct, you should say ‘DIAL’ and then press the “End Call” button. When you say ‘DIAL’ we will know that you are done with the task.

If the phone number is not correct, you should say ‘CLEAR’ to erase the incorrect number. The system will then tell you to say a phone number. At this point, you should repeat all of the digits, one at a time. They will still be visible on the Task Screen. If they are still incorrect, you should just go ahead and say ‘DIAL’ aloud and press the “End Call” button to make sure that you have a few seconds to prepare for the next phone number to be presented. In other words, after two attempts the task will be considered complete.

Then, you will wait for the next phone number to be presented. After you hear the new phone number and the word ‘Begin’, you will perform the same sequence, starting with the “Push to Talk” button. You will continue in this way until the trial is complete.

Do you have any questions or need a repeat of any instructions before we practice this task?
[Load PhoneTrain.tsv for stationary practice.]

AUDITORY-VOCAL ADDRESS ENTRY INSTRUCTIONS (4)

In this task you will enter an address into the navigation system by using voice commands. During each trial, you will enter several different addresses, one at a time. The addresses will be presented auditorily and will be visible on the Task Screen.

> **[All Participants**] In the driving simulator, the first address will be presented shortly after the lead vehicle appears.
> **[Non-Driving Participants Only**] In the non-driving venue, we will begin the trial when you are ready.

At the beginning of a trial, you will hear the first address followed by the word “Begin.” At this point, you should work quickly and accurately to complete the task using voice commands.

First, press the “Push to Talk” button on the right-hand side of the steering wheel, and listen for the system prompt to indicate it is listening for your voice command.

When the system is listening, you should say ‘Find an Address’, and then wait for the system prompt that will ask you to say the address. Then, say the entire address, starting with the house number, the street name, and the city name at a normal pace without pausing [all addresses are in Ohio]. If you pause, the system may get confused and not understand what you say after you pause.

When you have completed entering the address, listen for the system prompt that will confirm your entry and ask for the next command. The map screen should appear with the address displayed at the top. When the map screen appears, you should check to see that the address is correct. If it is correct, you should say ‘DONE’ aloud to indicate that the task is complete and press the ‘Home’ button to get ready for the next address.

If the address is not correct, you should press the ‘Home’ button and repeat the address entry.
This is done by pressing the “Push to Talk” button and saying ‘Find an Address’ when the system is listening. As before, the system will prompt you for an address in Ohio. You should say the address again. It will be available on the Task Screen.

When the map screen appears, you should check the address. If the address is still incorrect, you should just go ahead and say ‘DONE’ aloud and press the ‘Home’ button to make sure that you have a few seconds to prepare for the next address to be presented. In other words, after two attempts the task will be considered complete.

After you hear the new address and the word ‘Begin’, you will then perform the same sequence, starting with the “Push to Talk” button. You will continue in this way until the trial is complete.
Do you have any questions or need a repeat of any instructions before we practice this task?
[Load AddressTrain.tsv for stationary practice.]

N-BACK TASK INSTRUCTIONS (5)

The “N-back” is an auditory memory task. In this task, you will hear a voice recording of a sequence of numbers presented one at a time with several seconds between each number. After each group of 10 digits, there will be a brief pause, followed by the word ‘Next’. This sequence continues for the entire trial; the pauses provide a momentary break from the task. Your task will be to remember the most recent numbers and say a specified number aloud after each presentation. The specified number will be either “1-back” or “2-back” from the most recently presented number.

An example of the sequence of numbers you will hear is presented in the left-most column in each table below. First you will hear ‘4’ then ‘6’ then ‘7’ and so on. The responses that you should say aloud for the 1-back and 2-back conditions are presented in the right-most column of each table. Notice that in each condition, the sequence that you are to say aloud is the same as the original sequence. It is just delayed in both the 1-back and 2-back conditions. Example Sequence:

1-Back Task	
What You Hear	What You Should Say
Next	nothing
4	nothing
6	4
7	6
3	7
1	3
2	1
9	2
5	9
8	5
0	8

2-Back Task	
What You Hear	What You Should Say
Next	nothing
4	nothing
6	nothing
7	4
3	6
1	7
2	3
9	1
5	2
8	9
0	5

Let’s look more closely at the 1-back condition: After the word ‘Next’, the first number you will hear is ‘4.’ Because there is no 1-back number at this point, you will not say anything. Next you will hear ‘6.’ At this point you will say ‘4’ because this number is one back from the current number 6. Next, you will hear ‘7’ and you should say ‘6’ because it is one back from the current number 7. With the exception of the first number occurring after the word ‘Next’, you will say a number aloud immediately after hearing each number in the 1-back condition. You should say the number quickly so that you don’t miss the next number, which will be presented within a couple seconds. Now look at the 2-back sequence to see how it differs from the 1-back sequence.

You will be given instructions before each trial about which version of the task you will perform (1-back or 2-back). Most people will make mistakes in each condition. If you become aware that you have made a mistake, it might help to clear your memory by not responding to the item and effectively start over. Your performance score will be determined by the number of correct responses you make.

Do you have any questions before we practice this task? [Run 1Train.txt, then 2Train.txt for stationary practice.]

HEART RATE MONITOR INSTRUCTIONS

Now we are ready to attach the heart rate sensors. We will use two sensors. One will be attached just below the collarbone on your right side; the other will be attached near your lowest rib on the left. The two sensors will be connected to a lightweight plastic harness that will be attached to your shoulder.

[Show pictures]

We will use an alcohol pad to clean the spots on which we will attach the sensors. This helps us get a better reading from the EKG sensor, which measures your heart rate. If you are not comfortable having us do this, we will allow you to clean the spots and attach the sensors yourself. Or if you would feel more comfortable having a female /male research assistant attach these sensors, this can be arranged. If you have any questions about the sensor attachment, please feel free to ask me at any point. How would you like to attach the sensors?

> [If participant wants to apply sensors, give them materials and direct them to private location.]
>
> [If participant wants an experimenter of another gender] I will go now to get a male/female experimenter. I will be back in a few moments.
>
> [If experimenter is given permission to proceed]

I'm going to wipe your skin at the locations I described earlier with an alcohol pad before placing the EKG sensors. Would you raise your shirt/blouse on the left side a little for me? Thanks. [Use an alcohol pad to clean the participant's skin on the left side at approximately the level of the bottom rib. Then clean the area by the right collar bone.]

Now I am going to attach the EKG sensors. These monitor heart rate.

[Experimenter: Attach snap leads to the EKG electrodes before removing the electrodes from the plastic sheet. Run the leads out through the participant's shirt collar. EKG Sensors (2 contacts) – Active leads just under the right collar bone and over the bottom rib on the left side of the body create a vector across the heart. Clean skin with alcohol pad and wipe dry before placing sensors. EKG on right side; orient cable up over right shoulder. EKG placement over lower rib on the left side, exact placement is not critical for this lead. Note that the lead wires are brought together at the top of the right shoulder and then routed back toward the instrument plug-ins (placed on back side of right shoulder).]

Appendix E Simulator Sickness Questionnaire

Participant Number: ___

Directions:

Circle one option for each symptom to indicate whether that symptom applies to you right now.

1. General Discomfort...............None..............Slight..............Moderate ....Severe
2. Fatigue ....None..............Slight..............Moderate ....Severe
3. Headache ....None..............Slight..............Moderate ....Severe
4. Eye Strain ....None..............Slight..............Moderate ....Severe
5. Difficulty Focusing ....None..............Slight..............Moderate ....Severe
6. Salivation Increased ....None..............Slight..............Moderate ....Severe
7. Sweating ....None..............Slight..............Moderate ....Severe
8. Nausea ....None..............Slight..............Moderate ....Severe
9. Difficulty Concentrating ....None..............Slight..............Moderate ....Severe
10. "Fullness of the Head" ....None..............Slight..............Moderate ....Severe
11. Blurred Vision ....None..............Slight..............Moderate ....Severe
12. Dizziness with Eyes Open ......None..............Slight..............Moderate ....Severe
13. Dizziness with Eyes Closed ....None..............Slight..............Moderate ....Severe
14. *Vertigo ....None..............Slight..............Moderate ....Severe
15. **Stomach Awareness ....None..............Slight..............Moderate ....Severe
16. Burping.................................No..................Yes.................If yes, no. of times ______
17. Vomiting...............................No..................Yes.................If yes, no. of times ______
18. Other _____________________________________

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Appendix F Extra Information Pertaining to Experimental Procedures

Regarding Experimental Procedures:

The process of conducting experiments using human subjects (under Federal sponsorship) includes determination of participant compensation amounts and initiating applications for approval from two government agencies and one external entity:

1. Participant payment rate is determined using a method developed in conjunction with NHTSA's Office of Acquisition Management and is based on General Schedule pay level for comparable labor.

2. The Office of Management and Budget provides approval to collect information from U.S. citizens who may participate in the research including contact information and driving experience.

3. Appropriate agency personnel review and approve of means to protect and secure the information collected from both potential and selected participants.

4. An independent Institutional Review Board provides approval of the Informed Consent Form and the Experimental Protocol to ensure the safety of participants.

The approvals can take a significant amount of time, leaving a window of time to construct apparatus; to develop, test, and modify test scenarios; and to conduct pilot testing with internal employees who volunteer and participate anonymously.

Appendix G Complex Speed Car-Following Condition

The following (NADS ISAT) expression created the profile that was used to specify the lead-vehicle speed for the complex condition:

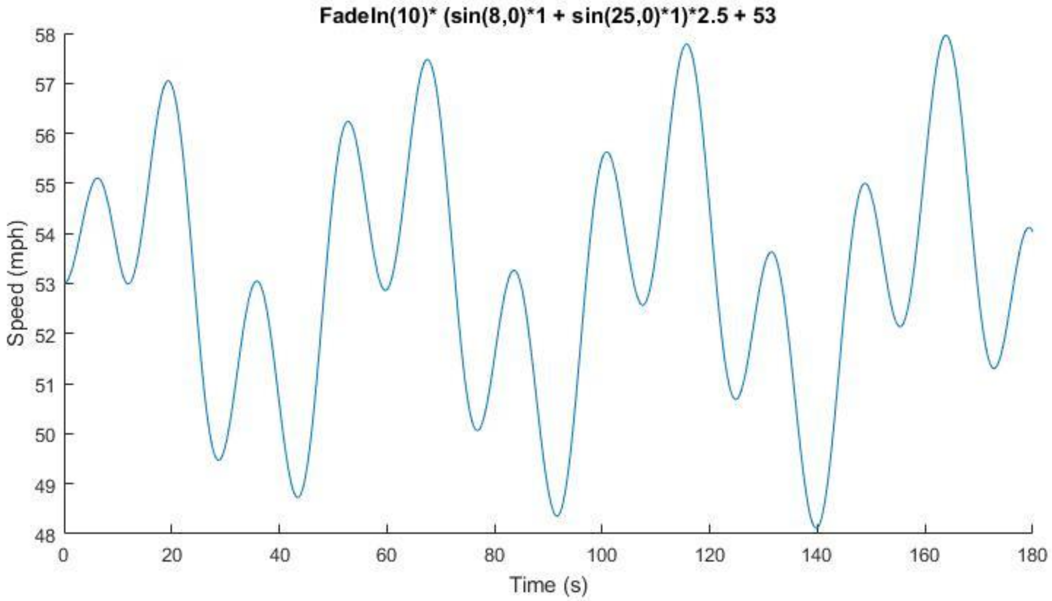
FadeIn(10) * (sin(8,0)*1 + sin(25,0)*1)*2.5 + 53

The initial speed is 53 mph. The FadeIn function [FadeIn(period)] is a simple ramp function that goes from 0 to 1 in a linear manner in the time given by period. When the time since activation exceeds the passed-in value period (time is seconds), FadeIn will return 1. FadeIn(10) goes from 0 to 1 in 10 seconds.

The sin function [sin(float half_period, float phase_offset)] returns a value based on the following equation:

sin((timeSinceStart + phase_offset)* 3.141592654 * 1/half_period)

The timeSinceStart refers to the time in seconds from the beginning of the car-following episode. The lead-vehicle speed signal based on this expression is presented in the following figure.



DOT HS 812 800
September 2019

U.S. Department
of Transportation

**National Highway
Traffic Safety
Administration**

NHTSA

12388-082019-v2