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

www.nhtsa.gov

DOT HS 811 463

November 2011

Developing a Test to Measure Distraction Potential of In-Vehicle Information System Tasks in Production Vehicles

DISCLAIMER

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT HS 811 463		

4. Title and Subtitle
Developing a Test to Measure Distraction Potential of In-Vehicle Information System Tasks in Production Vehicles

5. Report Date
November 2011

6. Performing Organization Code
NHTSA/NVS-312

7. Author(s)
Thomas A. Ranney, G. H. Scott Baldwin, Ed Parmer, and Joshua Domeyer, Transportation Research Center Inc.; John Martin, Ohio State University; 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 Avenue, SE.
Washington, DC 20590

13. Type of Report and Period Covered
Final Report

14. Sponsoring Agency Code

15. Supplementary Notes
The authors wish to thank Mike Perel, Adam Andrella, Jodi Clark and Scott Vasko for their contributions to this work.

Three experiments were conducted to assess a test of distraction potential for in-vehicle information systems (IVIS) and portable devices used while driving. The test used a low-fidelity (PC-based) driving simulator; sensors record drivers' control inputs in stationary production vehicles. Participants performed car-following and target detection together with secondary tasks. Experiment 1 examined the effects of two levels of driving task (car-following) difficulty and two detection tasks on test sensitivity. Detection tasks included a head-mounted task (HDT) and a computer-generated multiple-target task (MDT), which incorporated simple targets into the simulated roadway display. Secondary tasks included simple (Circles) and complex (navigation destination entry) visual-manual tasks and a hands-free auditory-vocal task (N-back). The MDT was more sensitive to task load differences, while the HDT created problems for the eye tracker. Increasing car-following task difficulty had no effect on metric sensitivity. The complex visual-manual task was more disruptive than the simple visual-manual task or the auditory-vocal task. The second experiment compared metrics provided by two occlusion paradigms: (1) traditional occlusion, which involves intermittent masking of the task to simulate the visual demands of driving, and (2) Enhanced Occlusion Task (EOT), which added auditory tracking to more realistically simulate driving task demands. Task duration estimates with EOT were closer to static completion times than those obtained with traditional occlusion; however their usefulness in estimating task duration requires a stronger connection to comparable estimates obtained from a driving protocol. EOT did not improve the R (task resumability) metric, although there was no independent evidence to support the expectation of differences in this metric. The R metric was not related to driving performance degradation. Auditory tracking performance metrics obtained in the EOT paradigm revealed performance degradation consistent with effects observed with simulator metrics. Experiment 3 used the simulator-based test to assess the distraction potential of navigation tasks performed with three systems with comparable functionality, including one original equipment manufacturer (OEM) and two portable systems, which had been independently rated as having different levels of usability. Metrics revealed strong and consistent differences between driving alone and driving with a secondary task. Three metrics (car-following coherence, detection task response time, and proportion of long glances) revealed differences between the (simple and complex) navigation tasks across all systems. Two metrics (standard deviation of lane position [SDLP] and detection task proportion correct) exhibited significant Systems x Task interactions; differences between navigation tasks were not consistent across systems. It was concluded that developing a simulator-based distraction potential test is feasible. Core metrics include those sensitive to visual-manual task conditions (SDLP, car-following delay, and detection task response time) and those sensitive to auditory-vocal task conditions (car-following delay, detection task response time, and detection task proportion of correct responses). Measures based on eye position data, primarily the proportion of long glances away from the forward roadway, revealed significant promise. Estimates of distraction potential can be combined with task duration estimates provided by the EOT to compute estimates of drivers' exposure to risk.

17. Key Words
Driver performance, distraction, cognitive distraction, navigation system, in-vehicle information systems (IVIS)

18. Distribution Statement
Document is available to the public from the National Technical Information Service www.ntis.gov

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	103	

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF TERMS

ADAM – Advanced Driver Attention Metrics – A joint initiative between DaimlerChrysler and BMW to gain a better understanding of the attentional demands of in-vehicle technologies and activities.

ADD – Destination entry by address – In-vehicle task performed using navigation systems, in which the user enters city, street name, and street number to obtain directions to a specified address.

AMD – Absolute Mean Deviation – Measure of auditory tracking error in the Enhanced Occlusion Paradigm. Higher values indicate poorer tracking performance.

ANOVA – Analysis of Variance

BAS – Baseline driving – Driving with no secondary tasks

CAMP – Collision Avoidance Metrics Partnership – Consortium of automobile manufacturers, including Ford and General Motors, formed in 1995 to facilitate cooperative pre-competitive industry/government research designed to accelerate the implementation of crash avoidance countermeasures.

CC – Cross-correlation – Computational approach used to develop alternate measures of car-following performance that correspond to coherence and phase shift, which are computed using frequency analyses.

Cir – Circles task **–** Reference/calibration task, developed as part of the HASTE program. This visual-manual search task requires participants to find a single (larger) target circle among a display of smaller targets. This task allows systematic increase of processing load by varying the relative sizes of the target and distractor circles.

COG – Cognitive **–** Secondary task with no visual or manual components. N-back is a cognitive secondary task used in the present work.

Cognitive distraction – Distraction occurs when the driver's attention is temporarily diverted away from the driving task while performing a secondary task. Cognitive distraction refers to the diversion of attention resulting from mental workload associated with tasks that involve thinking or memory.

Coherence (Cohere (Freq) or Cohere (CC)) – Coherence is a measure of car-following performance. Cohere (Freq) refers to the traditional use of frequency analysis to compute coherence. Cohere (CC) refers to the alternate use of cross-correlation to compute a comparable measure of car-following performance.

DT – Detection task – Generic reference to a target detection task used to assess driver workload, including the Peripheral Detection Task (PDT) and newer alternatives, including the Head-mounted detection task (HDT) and the Multiple target detection task (MDT).

DT MRT – Detection task mean response time – Detection task performance metric.

DT P Corr – Detection task mean proportion of correct responses – Detection task performance metric.

DWM – Driver Workload Metrics - The CAMP Driver Workload Metrics Project (2001-2005) brought together Ford, General Motors, Nissan and Toyota with the U.S. Department of Transportation to develop performance metrics and test procedures to evaluate the visual and cognitive aspects of driver workload from in-vehicle systems.

ENT – Enter button of navigation system

EORT – Eyes off road time – A measure of the total amount of time that the driver's eyes are diverted away from the forward roadway view, typically used to estimate the total amount of time required to complete a secondary task (i.e., task duration). EORT can be obtained either directly from eye glance data recorded during driving or from the sum of the shutter open intervals in the occlusion paradigm.

EOT – Enhanced Occlusion Task – combines traditional occlusion with an auditory tracking task to improve the validity of the obtained metrics by providing a task load more consistent with driving than occlusion alone.

FaceLAB – Eye tracking system developed by Seeing Machines, which uses unobtrusive cameras to record and compute the position of the driver's head and eye gaze.

GPS – Global Positioning System – Portable navigation systems are referred to as GPS devices.

HASTE – Human machine interface And the Safety of Traffic in Europe – Eight European partners and Transport Canada conducted this project, which was intended to develop methodologies and guidelines for the assessment of In-Vehicle Information Systems (IVIS).

HDT – Head-mounted Detection Task- variation of Peripheral Detection Task in which a single target is affixed to the head. This task keeps the position of the target constant relative to the driver's eye position.

Hdwy – Headway – distance between lead and following vehicle in car-following

ISO – International Standards Organization

IVIS – In-Vehicle Information Systems – Navigation systems are examples of IVIS.

LCD – Liquid Crystal Display

LED – Light-Emitting Diode

LS – List Search – Navigation system task, in which drivers search a list of destinations selected by the system.

LV – Lead vehicle in a car-following task. Drivers are instructed to maintain a constant following distance.

MDT – Multiple target Detection Task – Alternative to the Peripheral Detection Task (PDT) in which targets are presented in the simulated roadway display at different locations.

MicroDAS – Data acquisition system developed by researchers at NHTSA's Vehicle Research and Test Center (VRTC).

MRT – Mean Response Time – Detection task performance measure.

NAV – Navigation System – Both OEM and portable GPS devices are navigation systems.

N-back – Artificial secondary task, in which participants are required to listen to and recall a stream of digits. N refers to the position in the stream that the participant is required to recall. As N increases, the demands of the task increase, i.e., 2-back is more demanding than 1-back.

NHTSA – National Highway Traffic Safety Administration

Occlusion – An experimental technique in which tasks are intermittently masked, typically using occlusion goggles, which alternate between transparent and opaque states. The masked intervals are intended to simulate the real-world requirement of looking away from the secondary task to monitor driving conditions.

OEM – Original Equipment Manufacturer – A factory-installed navigation system is referred to as an OEM system.

P Corr – Proportion of correct responses – Target detection performance measure.

PC-based – Personal computer-based – Both the driving simulator and the occlusion tasks used in the present work were implemented on PCs.

PD – Previous Destination – Secondary task performed using a navigation system in which the participant must locate and select a destination that has previously been entered in the system's memory.

PDT – Peripheral Detection Task – Simple target detection task, typically performed together with driving tasks to assess driver workload. Traditional implementation involves periodic illumination of one of several LEDs, which is reflected from vehicle windshield. Drivers must respond as quickly as possible to the onset of LEDs by pressing a button attached to their finger.

PLATO – Portable Liquid Crystal Apparatus for Tachistoscopic Occlusion – Technology implemented via goggles used in the occlusion protocol. PLATO goggles allow for periodic interruption of the participant's vision to simulate the visual demands of a driving situation in which the driver's visual attention is switched between the road ahead and a secondary task inside the vehicle.

P RAP – Percentage of time inside region of acceptable performance – Measure of auditory tracking performance from the Enhanced Occlusion Task paradigm.

PRC – Percent Road Center – A metric derived from eye position data recorded by an eye tracker that represents the percentage of time during a driving trial that a driver is attending to the forward roadway. Higher PRC values indicate more attention to the roadway ahead, while lower PRC values indicate diversion of visual attention away from the roadway ahead, most often due to a secondary task competing for the driver's visual attention.

R – Resumability metric – An attribute of a secondary task, which is computed from measures obtained in the occlusion protocol. It represents the relative ease with which a particular task can be completed under conditions of interrupted performance, as in performing the task while driving.

RAP – Region of Acceptable Performance – In the Enhanced Occlusion Task protocol, this metric characterizes auditory tracking performance. It refers to a predefined region on either side of the target signal, such that tracking performance is considered to be error-free while the participant maintains the cursor within this region.

RSME – Rating Scale Mental Effort – A single scale used to record participants' subjective assessment of the mental workload or effort associated with a given task condition. It is used to assess subjective workload both in driving simulator and occlusion experiments.

RT – Response Time – Measure of detection task performance.

SAS – Statistical Analysis Software – Commercial product, widely used for data analysis.

SDLP – Standard deviation of lane position – Driving performance metric that characterizes the lateral movement of the vehicle. Larger values of SDLP have been interpreted as evidence of inattention to steering control, typically associated with secondary tasks that require removal of drivers' hands from the steering wheel.

SRD – System Response Delay – IVIS systems, particularly those that require satellite communication, are subject to delays between the time a user makes an input and when the system responds. These delays, if significant, can influence the task duration estimates obtained with occlusion protocols.

Std – Standard Deviation – A summary measure that characterizes the amount of variation in a particular sample of data.

Std Hdwy – Standard deviation of headway – This metric characterizes the amount of variation inherent in a sample of headway data obtained from a single driving trial. Drivers are instructed to maintain a constant following distance (headway), therefore larger values of Std Hdwy typically reflect increased inattention to the car-following task.

STI – Systems Technology Incorporated – developer of the STISM-Drive Simulator.

STISIM – Systems Technology Incorporated Driving Simulator – PC-based simulator also referred to as STISIM-Drive.

SV – Subject Vehicle – Simulated vehicle controlled by experimental participant while performing secondary task.

TNO – Netherlands Organization for Applied Scientific Research – cited in this work for developing the HDT.

TRC Inc. **–** Transportation Research Center Inc. **–** Organization that performed current research under contract to NHTSA.

TSOT – Total Shutter Open Time – Metric derived from the occlusion paradigm, which estimates the total time required to complete a task based on the total duration of all intervals in which the task was visible to the driver.

TTT – Total Task Time **–** Metric derived as part of occlusion protocol that characterizes the total amount of time required to complete a task. It is measured in a static situation, involving continuous, uninterrupted performance.

UNOCC – Unoccluded – Refers to the time in the occlusion protocol in which the vision is not occluded. Total Shutter Open Time (TSOT) is the sum of all unoccluded intervals.

UTC – Coordinated Universal Time

VRTC – NHTSA's Vehicle Research and Test Center, located in East Liberty Ohio.

EXECUTIVE SUMMARY

The measurement of distraction has been the focus of several large-scale projects undertaken by consortia of researchers, government agencies, and automotive manufacturers in recent years. This work has been directed at the need to evaluate pre-production versions of in-vehicle systems, sometimes referred to as in-vehicle information systems (IVIS). Recently, researchers at the National Highway Traffic Safety Administration's (NHTSA) Vehicle Research and Test Center (VRTC) undertook a study to assess the feasibility of adapting protocols and measures to assess distraction associated with driver interaction with in-vehicle systems that are already available in production vehicles. Based on an evaluation of the most promising distraction metrics, a prototype test was developed, which included the combination of car following and peripheral target detection. The test was implemented on the STISIM, which is a low-fidelity (PC-based) driving simulator. Target detection was implemented using the Peripheral Detection Task (PDT), which requires rapid responses to simple and frequently-occurring targets that appear in the driver's peripheral visual field. The metrics selected for further development included measures of driving performance (car-following delay, lane-position variability, and steering error) and visual target-detection performance (mean response time, proportion correctly detected). As part of this previous work, three issues were identified that required additional development. The first issue concerned the need for additional sensitivity for detecting performance degradation due to cognitive distraction, which refers to the diversion of the driver's attention away from driving as the result of mental activities, such as thinking, remembering, or evaluating options. The second issue concerned the need to develop a method to obtain steering inputs from production vehicles with minimal setup time and without damage to the vehicle. The third issue concerned the need for improving the quality of the eye position data to support the computation of eye-glance metrics. The first objective of the present work was to develop and evaluate solutions to these methodological problems.

The second objective of the present work was to determine whether the occlusion technique, and in particular an enhanced version of this technique, provided information that could help in the assessment of the distraction effects of in-vehicle secondary tasks. The occlusion technique involves periodic interruption of vision (via occlusion goggles) during the performance of a secondary task (e.g., navigation system destination entry); it provides an estimate of the time that the driver must look away from the roadway to perform a particular secondary task. Data from occlusion trials are also used to compute indices of task resumability (R), which indicate how amenable a task is to completion under conditions of interruption, as in driving. The Enhanced Occlusion Technique (EOT) combines the traditional occlusion technique with a computer-generated auditory tracking task. It was developed to improve the validity of task completion time estimates.

The third objective of the present work was to incorporate improvements to the test protocol and use the revised protocol to assess the distraction potential of multiple systems, including an original equipment manufacturer (OEM)-installed navigation system and portable devices with comparable functionality. The objectives were addressed in three experiments.

This research was conducted in 2009. All three experiments were conducted in stationary vehicles, which were not running. For experiments 1 and 3, the vehicles were equipped with steering, brake, and throttle sensors to provide control inputs to the driving simulator. Drivers performed the secondary tasks while driving the simulator with the stationary vehicle controls.

Changes in driving performance measures in this dual-task condition relative to a baseline condition, which involves driving only, were interpreted as an indication of the amount of distraction potential associated with the secondary tasks. In the second experiment, participants performed secondary tasks under two occlusion conditions; one condition involved simple occlusion and one involved the EOT. Task completion times under different conditions were recorded and used to compute a measure of task resumability.

The first experiment evaluated two variants of the PDT and two levels of driving task demand, in an attempt to improve the sensitivity of metrics for detecting differences in cognitive distraction. Specifically, a head-mounted detection task (HDT) was compared with a computer-generated multiple-target detection task (MDT) that incorporated simple targets into the simulated roadway display at different locations. It was hypothesized that the head-mounted display task would provide better sensitivity for detecting effects of cognitive distraction by virtue of the fact that the target remained at a constant location relative to the driver's eyes. It was also hypothesized that increasing primary (driving) task demands, by increasing car-following task difficulty, would increase metric sensitivity.

Participants performed three categories of secondary tasks, including a simple visual-manual task (Circles), a complex visual-manual task (navigation system destination entry), and a hands-free auditory-vocal task (N-back), for which the distraction was primarily cognitive. The hypothesized difference between detection task conditions was not observed. To the contrary, the target-detection (response-time) metric was sensitive to differences between levels of the auditory-vocal task and the navigation tasks when the MDT was used but not when the HDT was used. Moreover, the HDT was associated with an unanticipated yet significant amount of deterioration of the eye position data; the head-mounted target apparently confused the eye tracker concerning the position of the driver's gaze while driving. The results also showed that increasing the car-following task demands did not significantly increase metric sensitivity.

More generally, it was found that most metrics were sensitive to differences between task categories; as expected, the complex visual-manual task was associated with greater driving performance degradation than the simple visual-manual task or the auditory-vocal task. Subjective ratings of mental workload, obtained using a single scale administered after each trial, were not entirely consistent with these results; participants generally rated the auditory-vocal task as more demanding than the simple visual-manual task. However, among objective measures, the simple visual-manual task was associated with higher levels of performance degradation for measures sensitive to visual-manual differences, while there were no differences between these tasks for measures more sensitive to cognitive differences. Driving performance metrics, including car following and detection task measures, were generally able to differentiate between conditions in the audio-vocal (N-back) task and in the navigation system tasks, but not for the Circles task, due to the relatively small difference in demand between conditions for this task.

The second experiment compared the traditional occlusion protocol with the EOT. Because participants have no primary task load (to simulate the demands of driving), the task-completion time estimates using traditional occlusion do not include time during which participants continue to work on the secondary task during occluded intervals. The EOT addresses this concern, called blind operation, by adding an auditory tracking task, intended to simulate the demands of driving without interfering with the visual demands of occlusion. The objectives of Experiment 2 were to determine the extent to which blind operation is eliminated by the EOT and to determine whether the EOT improves the sensitivity of the derived R metric, relative to the traditional occlusion protocol. Three navigation system tasks were used in Experiment 2, including destination entry by address (ADD), selecting a previous destination (PD), and searching a list of cities (LS).

The EOT eliminated part of the blind operation, but not all of it. Specifically, with traditional occlusion, approximately 23 percent of the effort required to perform the task was accomplished during occluded intervals. With the EOT, the corresponding percentage was 11 percent. The R metrics differed between the traditional occlusion and EOT conditions, but neither R metric revealed differences between secondary task conditions. This led to the conclusion that task resumability (R) does not reflect the same performance degradation revealed by the driving performance metrics. The ADD task was associated with a significantly higher level of (auditory) tracking error than the PD task, which is consistent with the simulator test results. This result implies that task duration estimates obtained with the occlusion technique must be considered together with the level of primary task degradation to provide a complete understanding of the effects of secondary tasks.

Experiment 3 incorporated modifications to the test protocol based on the results of Experiment 1, including use of the multiple-target detection task (MDT) and the moderate (less difficult) level of car-following task difficulty. The modified test protocol was used to assess the distraction potential of three navigation systems with comparable functionality. Participants performed two navigation system tasks (ADD and PD) using one OEM system and two portable systems, which differed in their rated usability. In a separate consumer product study, the High-Usability system was rated as easier to use than the Low-Usability system. It was hypothesized that the OEM product, by virtue of its design to be used specifically in the driving context, would be less potentially distracting than either of the portable systems. Based on the assumption that usability ratings are correlated with the potential for distraction, it was also hypothesized that the High-Usability system would be less potentially distracting than the Low-Usability system. Metrics revealed strong and consistent differences between baseline driving and driving with a secondary task. Three objective metrics (Car-following coherence, detection task mean response time and the proportion of long glances) revealed differences between the ADD and PD tasks generally; however these differences were weaker than those observed in Experiment 1, reflecting the fact that the patterns of results for the three systems were not consistent. Specifically, the SDLP metric exhibited a significant interaction between Systems and Tasks. As predicted, the ADD task was more distracting than the PD task for the OEM and High-Usability systems, but contrary to predictions, the reverse was true for the Low-Usability system. A similar pattern was observed for the detection task proportion of correct responses. The occurrence of complex interactions indicates that there may be subtle differences between systems that affect the potential for distraction and that conclusions about the distraction potential of a particular task cannot be made without considering the system on which the task was performed. It also suggests that usability ratings may not be highly correlated with distraction potential for some devices.

Based on the results of these experiments, we concluded that the development of a simulator-based test to assess the distraction potential of secondary tasks performed with OEM equipment in production vehicles or portable devices is feasible. The test can be implemented without requiring significant setup and without damaging vehicles. The test focuses on the dynamics of distraction and does not consider the duration of the distracting activity, which is necessary to fully characterize the exposure to risk associated with a distracting activity. The EOT represents an improvement over the traditional occlusion paradigm for providing information about the time required to perform various secondary tasks; however, task duration estimates obtained with the traditional occlusion protocol or the EOT require a stronger connection to comparable values obtained in a controlled driving situation.

Test results indicated that a broad range of metrics, including measures of car-following, lateral vehicle control, target-detection, and visual performance, were consistently and robustly sensitive to differences between categories of secondary tasks and between baseline driving and driving while performing secondary tasks. Fewer metrics were found to be sensitive to differences between conditions within task categories. Metrics sensitive to differences between visual-manual task conditions included lane-position variability (SDLP), car-following delay and detection task response time. Metrics sensitive to differences between auditory-vocal task conditions included car-following delay, detection task response time, and detection task proportion of correct responses.

Due to their increased sensitivity for detecting differences within task conditions, the SDLP, car-following delay, detection task response time and proportion of correct responses are considered core metrics for assessing distraction potential. Measures based on eye position data, primarily the proportion of long glances away from the forward roadway, exhibited differences between conditions within tasks, but the results were weaker and less consistent than the differences observed for performance-based metrics.

Establishing levels of acceptable dose, particularly for cognitive distraction, remains a significant challenge. The N-back task provided a significant dose of cognitive distraction that was consistently disruptive to driving performance. Based on the present results, the 2-back condition could serve as a starting point for defining a limit for acceptable “dose” of cognitive distraction.

1.0 INTRODUCTION

1.1 Background

The measurement of distraction has been the focus of several large-scale projects undertaken by consortia of researchers, government agencies, and automotive manufacturers. These include the European project HASTE (Human machine interface And the Safety of Traffic in Europe) (Carsten & Brookhuis, 2005; Carsten et al., 2005), the Driver Workload Metrics (DWM) Consortium of the Collision Avoidance Metrics Partnership (CAMP) (Angell et al., 2005) and the German Advanced Driver Attention Metrics (ADAM) program (Mattes, 2003). The goal of these projects has been to develop methodologies and guidelines for assessing the extent to which in-vehicle information systems (IVIS) interfere with driving. The Alliance of Automobile Manufacturers has developed guidelines based on this work. The ISO (ISO, 2004; ISO 2007) continues to work on developing standard procedures for measuring driver workload.

Much of this work has been directed at the original equipment manufacturers' (OEM) need to evaluate pre-production versions of IVIS, thus allowing design modifications if necessary before a vehicle is released. As a result, not much consideration has been given to adapting protocols or measures to assess IVIS that are already available in production vehicles. The National Highway Traffic Safety Administration (NHTSA) anticipated the need to assess IVIS in production vehicles to assess compliance with and/or to establish guidelines and undertook a project, conducted by researchers at NHTSA's Vehicle Research and Test Center (VRTC), to adapt one or more existing protocols for this purpose. Three experiments were conducted to evaluate the most promising metrics. Based on the results of these experiments, the combination of car following and peripheral target detection was selected for further development and evaluation. Following CAMP, the car-following task was implemented on the STISIM, which is a low-fidelity (PC-based) driving simulator. Target detection was implemented using the Peripheral Detection Task (PDT), a dashboard-mounted array of LCDs that create reflections in the windshield in the peripheral visual field. The PDT has been used in numerous studies (e.g., Harms & Patten, 2003) both as a measure of workload and as a measure of object and event detection, a component of driving behavior. The combination of car following and target detection offers significant flexibility for fine-tuning scenario components plus a wide range of performance measures.

Results of the initial tests of this car-following/PDT combination test venue were presented in a recent report (Ranney, Baldwin, Vasko, & Mazzae, 2009). In addition to the specific metrics selected for further development, two issues were identified that require additional development. The first issue concerns the need for additional sensitivity for detecting performance degradation due to cognitive distraction, which refers to the diversion of drivers' attention away from driving due to tasks that are primarily mental and have no visual-manual components. Thus, while the metrics were generally found to be sufficiently sensitive for distinguishing between different levels of demand associated with IVIS tasks performed with visual-manual interfaces, they were less sensitive to such differences associated with tasks that use auditory interfaces, in which the distraction is primarily cognitive.

We identified two strategies for improving the sensitivity for detecting effects of cognitive distraction. The first strategy involved modifying the traditional peripheral detection task (PDT) based on emerging research results, which have suggested that newer variations can provide at least comparable sensitivity for detecting cognitive distraction, while at the same time providing greater operational flexibility. Specifically, Victor and colleagues (Victor, Engstrom, & Harbluk, 2008) have argued that a detection task consisting of a single centrally-located target or a single head-mounted target would provide greater sensitivity than the traditional PDT, which uses an array of targets displayed in the peripheral visual field. Accordingly, one objective of this project was to evaluate several alternative detection tasks to determine which provides the most sensitivity for detecting the effects of cognitive distraction. The second strategy for increasing metric sensitivity to effects of cognitive distraction involved increasing the primary (driving) task demands to reduce the amount of spare attentional capacity available to perform secondary tasks. As primary task demands are increased and spare capacity is reduced, we expect to observe primary task degradation at relatively lower levels of secondary task demand, reflecting increased metric sensitivity. We evaluated this hypothesis in this work.

The second issue was a problem of measurement; it concerned the way in which steering inputs are obtained from production vehicles. In the previous study, we used an overlay steering wheel, which allowed us to obtain steering inputs without requiring that the vehicle be running (to activate power steering). The use of the overlay steering wheel was acceptable for obtaining steering inputs but created problems when IVIS systems required use of buttons located on the vehicle steering wheel, which were not readily accessible due to the overlay. We developed and evaluated two engineering solutions as part of this study. One approach involved rotating plates connected to the vehicle's front tires such that the tire rotation was recorded and used to measure steering inputs. The second approach used gravity-based inclinometers attached to the steering wheel. Both approaches appeared suitable for use in our test protocol, which requires relatively quick installation on a wide range of vehicles.

1.2 Occlusion Technique

Measures of visual attention are emerging as strong indicators of distraction potential, reflecting the conclusion that the crash risk increases with the amount of time a driver looks away from the forward roadway. The occlusion technique, which involves periodic interruption (via visual occlusion) of the performance of an IVIS task (e.g., navigation system destination entry) (Stevens, Bygrave, Brook-Carter, & Luke, 2004), provides an estimate of the time that the driver must look away from the roadway to perform a particular secondary task (ISO, 2007). The periodic interruption of IVIS task performance is intended to simulate the real-world requirement of switching vision and attention between the IVIS task and driving. The Enhanced Occlusion Technique (EOT), which combines the traditional occlusion technique with a computer-generated auditory tracking task (Schindhelm & Gelau, 2008, 2009), was developed to address a methodological problem with the traditional occlusion technique. Specifically, with no processing load during the occluded intervals, the traditional technique allows participants to continue working on the IVIS task and thus does not provide a valid simulation of the disruption of IVIS task performance caused by the demands of driving. The resulting values of total shutter open time (TSOT) therefore do not include all of the time required to complete the IVIS tasks. Preliminary results using the EOT suggest that the time estimates may be more realistic estimates of the time required to perform the IVIS tasks while driving. EOT trials also provide information used to compute indices of task interruptability or resumability (R), which are indicative of how amenable a task is to completion under conditions of interruption. The addition of the EOT to our test protocol will allow us to determine whether the R values provide comparable sensitivity relative to the car-following/PDT for discriminating among different levels of IVIS task difficulty.

1.3 Study Overview

Three experiments were conducted in this 2009 study. The first experiment addressed the methodological problems identified in our previous work. Specifically, two detection task variants and two levels of driving task were evaluated. The objective was to determine whether the methodological modifications were associated with increased sensitivity, particularly for detecting differences in tasks in which the distraction effects were primarily cognitive. The results of this experiment were used to make improvements to the test protocol. The second experiment compared the traditional occlusion protocol with the EOT. The objective was to determine whether methodological modifications incorporated in the EOT improved the sensitivity of the occlusion metrics for differentiating among tasks with different task demands. The third experiment applied the modified (simulator/detection task) protocol to assess the distraction potential associated with a variety of IVIS tasks including navigation systems in a single production vehicle and two portable devices with comparable functionality. All three experiments were conducted in stationary vehicles, which were not running. For Experiments 1 and 3, the vehicles were connected to the STISIM simulator; steering, brake, and throttle sensors provided control inputs to the driving simulator. Drivers performed secondary tasks while driving the simulator (Experiments 1 & 3). Changes in driving performance measures in this dual-task condition relative to a baseline condition, which involved driving only, were interpreted as an indication of the amount of distraction potential associated with the secondary tasks. In the second experiment, participants performed secondary tasks under two occlusion conditions, including simple occlusion and the EOT. Task completion times under different conditions were recorded and used to compute a measure of task resumability.

2.0 EXPERIMENT 1

2.1 Background

The traditional PDT implementation uses a dashboard-mounted array of LCDs that create reflections in the windshield in the peripheral visual field. According to Victor et al. (Victor, Engstrom, & Harbluk, 2008), research using the PDT has failed to show the hypothesized effect of target eccentricity, according to which increased workload is associated with decreased sensitivity for detecting targets at increasing distance from the center of the visual field. This appears to be inconsistent with the finding that increased workload is associated with increasingly centralized gaze concentration, which has been found with measures of eye gaze position (Recarte & Nunes, 2003; Victor, Harbluk, & Engstrom, 2005). The failure to demonstrate an effect of target eccentricity with the original PDT led Victor and colleagues (Victor et al., 2008) to conclude that the “peripheral” aspect of the test was not valid. This conclusion has motivated recent demonstrations that alternative detection tasks, which involve presenting targets via different modalities (e.g., auditory, tactile), or in different locations (e.g., head-mounted) provide comparable sensitivity for detecting effects due to cognitive distraction (Victor et al., 2008).

It is possible, however, that technical limitations of the original PDT evaluations may have been at least partly responsible for the failure to find target eccentricity effects. Originally, the PDT was used primarily in real-world driving, which created significant difficulties with respect to the effect of sunlight on target brightness. In addition, the approach of using reflected targets limited the target location to a relatively small area of the peripheral visual field. The necessary freedom afforded to drivers to move their heads and eyes during driving also made it very difficult to control target eccentricity. Finally, the use of the PDT in real-world settings limits the experimenter’s ability to control the complexity and location of other information in the driver’s visual field, much of which required processing at a higher priority than PDT targets due to its potential safety relevance. Elimination of these problems could facilitate a more rigorous test of the attentional narrowing hypothesis, which if valid could provide a useful metric for assessing cognitive distraction.

One objective of the current work, therefore, is to evaluate two alternatives to the original PDT, both of which provide greater control of the target presentation. The first alternative is a head-mounted detection task; a single LED is attached to a headband such that the target always appears in the periphery of the driver’s visual field. This approach, which has been developed and used by TNO researchers, controls the target location relative to the driver’s eyes. The close proximity of the target to the driver’s eyes serves to control the target brightness. However, the use of a single target does not allow for evaluation of target eccentricity effects. The second alternative addresses this issue. Specifically, we evaluated a detection task in which targets are presented graphically on the simulator screen (Victor et al., 2008; Merat & Jamson, 2007). Screen presentation eliminates the target location constraints associated with the original dashboard-mounted task and allows presentation of targets at different eccentricities relative to the point at which drivers’ attention is assumed to be focused. The assumption of a fixed point of focus is based on the use of a car-following task in which drivers are required to maintain a consistent following distance from a lead vehicle and the corollary assumption that doing so will ensure that the driver’s attention is concentrated on the rear of the lead vehicle. Target locations were defined relative to this position.

The second objective of Experiment 1 was to evaluate the hypothesis that increasing primary (driving) task demand would increase metric sensitivity for detecting differences between different levels of demand, particularly for tasks performed with auditory interfaces, thus involving primarily cognitive distraction. This hypothesis is based on the assumption that when primary task demands are relatively low, drivers retain sufficient resources to concurrently perform a variety of secondary tasks of different loads, without any discernible effect on primary task performance. Measures of primary task performance are not sensitive to differences in secondary task load when these tasks can be performed within the limits of the driver's spare capacity. When primary demands increase, drivers must devote more resources to maintain acceptable primary task performance, thus reducing the spare attentional capacity available for secondary tasks with no discernible consequence. According to this conceptualization, reduced spare capacity will lead to degradation in primary task performance at relatively lower levels of secondary task demand, relative to the situation in which primary task demands are low and spare capacity is greater. Several of the HASTE studies presented findings that support the feasibility of this approach. Jamson and Merat (Jamson & Merat, 2005) manipulated primary task demand by comparing performance on straight (easy) versus curved (difficult) driving segments. They found improved differentiation among levels of secondary task load on measures of primary task performance at the more difficult level of primary task demand. Specifically, they found differences in lateral performance (increased lane position variability) for different levels of visual-manual secondary task demand on curved road segments that were not apparent on straight road segments. Briem and Hedman (Briem & Hedman, 1995) manipulated primary task demand by comparing performance on firm (easy) versus slippery (difficult) roads. They found increased sensitivity on primary task measures for detecting differences among secondary task conditions when the primary task demands were difficult. Specifically, they found differentiation among secondary task conditions on lane-position deviation in the slippery road condition but not in the firm road condition. These effects were more pronounced for secondary tasks that required physical manipulation than for those that required only hands-free communication.

In the present study we increased driving task demand by varying the difficulty of the car-following task. Specifically, we defined car-following difficulty in terms of the acceleration and deceleration requirements of the lead-vehicle speed signal. Accordingly, we assumed that when the driver was actively accelerating or decelerating, a higher level of conscious attention was directed to the car-following task than when traveling at a constant speed. We used a complex signal rather than a simple sinusoidal signal to increase the realism of the car-following task. The construction of the complex signal is described in an earlier study (Ranney, Mazzae, Baldwin, & Salaani, 2007).

2.2 Overview

The experimental objectives were addressed in an experiment in which participants performed different categories of secondary tasks while performing the combination of car following and target detection. Each participant performed under one of the four combinations of detection task and driving task difficulty. The experiment was conducted in a single stationary vehicle, which was connected to the driving simulator; steering, brake, and throttle sensors provided control inputs to the driving simulator. The vehicle was not running. Changes in driving performance measures in the dual-task condition relative to a baseline condition, which involves driving only, were interpreted as an indication of the amount of distraction potential associated with the secondary tasks.

2.3 Experimental Design

Experiment 1 used a mixed design, including both within- and between-subject factors. The main between-subject design factors (independent variables) were the target detection task (2 levels) and the driving task difficulty (2 levels). Each participant completed one of the four combinations of these factors, as shown in Table 1.

Table 1. Sample Size by Experimental Conditions (Experiment 1)

	Detection Task Condition	
Speed Profile	Head-Mounted	Multiple Targets
Moderate	10	10
Difficult	10	10

The Secondary Task condition (7 levels, including 2 levels of three tasks, plus a baseline condition) was varied within subjects, such that each subject completed all conditions. Thus, excluding training and practice, each participant completed seven three-minute drives. Simulator drives consisted of close car following on a straight road with minimal other traffic present.

2.3.1 Driving Task

A car-following paradigm modeled after that used by Brookhuis and colleagues (Brookhuis, Waard, & Mulder, 1994), was programmed into the scenario run on the STI simulator. This task required participants to maintain a constant following distance behind a lead vehicle, which changed speed according to a predefined complex waveform (see Figure 1). Participants were required to follow a simulated lead vehicle's speed changes on straight road segments. Prior to testing, drivers were given training and feedback about the range of following distances considered acceptable. During the experiment, participants received feedback and monetary incentives based on their ability to maintain an acceptable following distance. An auditory warning system was used to encourage drivers to maintain a fairly close following distance. When drivers exceeded a pre-defined criterion, an audible tone sounded once every five seconds until the driver returned to an acceptable following distance.

Figure 1 presents the variations of lead vehicle speed signal that were created for Experiment 1. The 'moderate' signal is the signal that had been used previously. The difficult signal was created by increasing the y-axis scaling of the moderate signal around its mean, which had the effect of retaining the same relative frequency components while increasing the amplitude. Car-following task difficulty was thus defined operationally as the standard deviation of acceleration at each point on the curve.

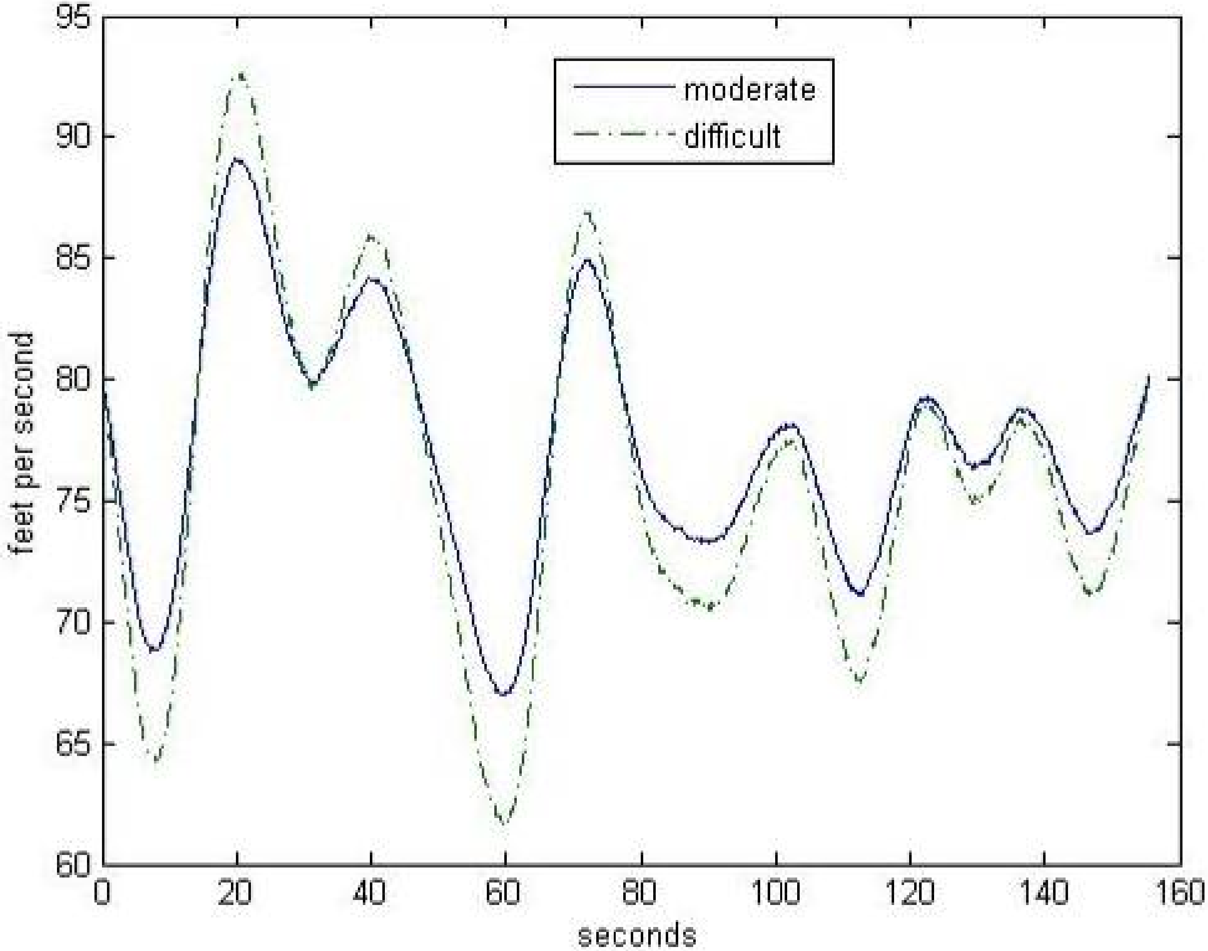


Figure 1. Lead Vehicle Car-Following Speed Signals – Experiment 1

2.3.2 Target Detection Tasks

Detection task variants included a computer-generated multiple target detection task (MDT) and a head-mounted task detection task (HDT). The HDT used a single LED, which was attached to a head-mounted apparatus shown in Figure 2. The apparatus was fabricated from a construction helmet suspension and weighs approximately 6 ounces. Mounting the LED on the head allowed the target to remain in the same position relative to the driver's eye position. In the MDT, targets (red-colored circles approximately the same size as the LED reflections in the traditional PDT) appeared at one of 6 locations on a single horizontal line near the horizon in the driving scene (see Figure 3). Thus, the two tasks differed in their frame of reference: the (HDT) head-mounted target appeared in the same position relative to the driver's head while the MDT targets appeared at fixed locations on the screen, which is more consistent with the traditional PDT.

Figure 2. Head-Mounted Detection Task (HDT)

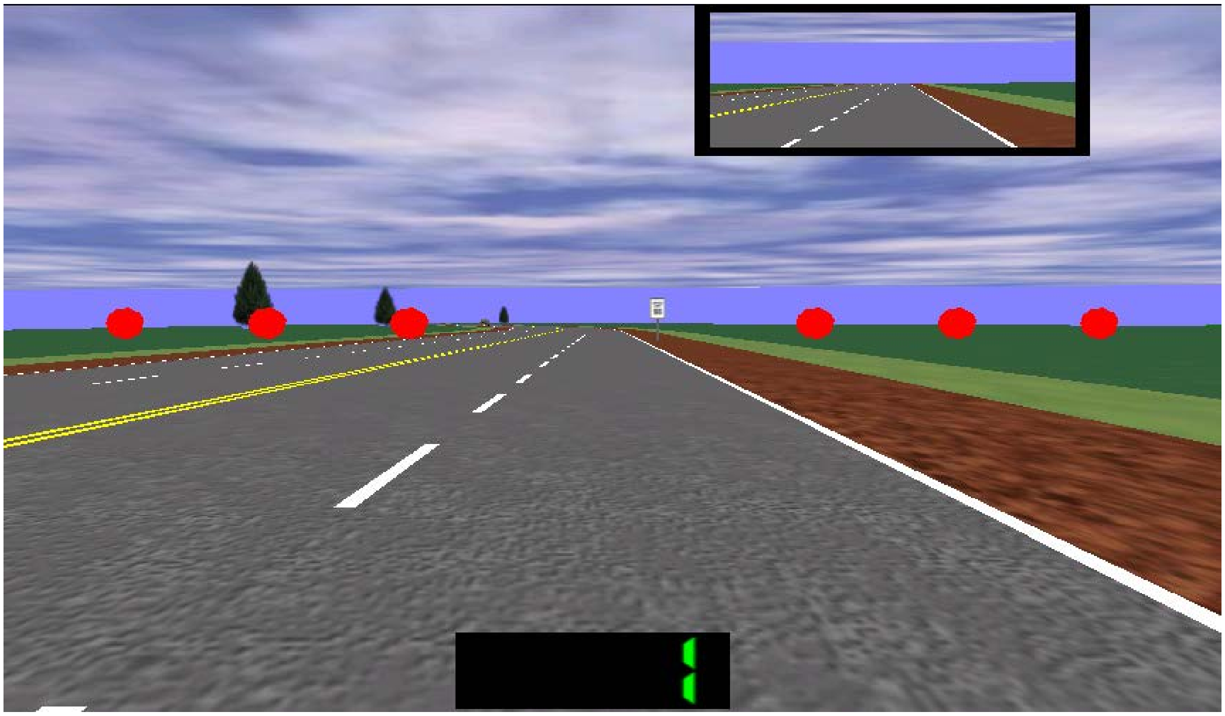
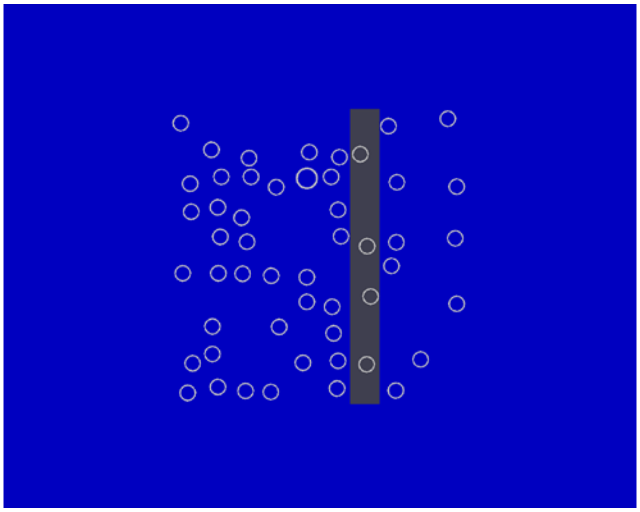


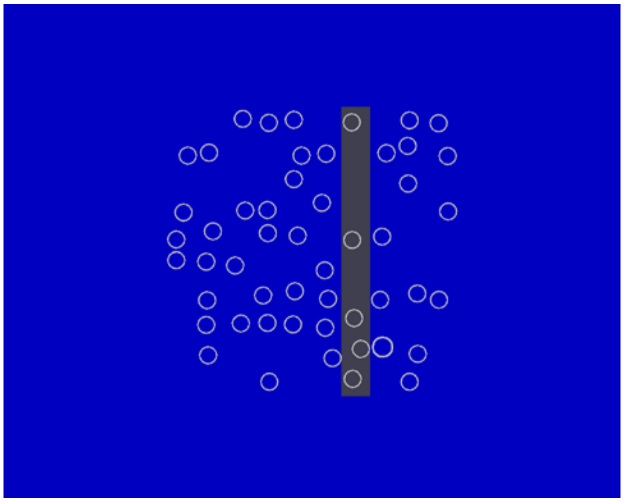
Figure 3. Multiple-Target Detection Task (MDT)

2.3.3 Secondary Tasks

Three categories of secondary tasks were used in the experiment, including: (1) the Circles task; (2) the N-back task, and (3) real-world navigation system tasks involving entering a new destination or selecting a previously-entered destination. The Circles task is a self-paced visual search task presented on a computer screen located inside the test vehicle, in which participants search an array of circles for a designated target. The target is a circle that is slightly larger than the other (distractor) circles. Participants respond by moving a vertical band on the computer screen (via button press) to the location of the target circle. The difficulty of the search task is manipulated by varying the relative sizes of the smaller (distractor) circles and the larger (target) circles. The difficulty of the response task can be varied by changing the size of the vertical band, which determines the number of button presses required to align the vertical band with the target location. Specific conditions used in this experiment are presented in Figure 4.



D115 (Distractor size: 115, Target size: 150)



D130 (Distractor size: 130, Target size: 150)

Figure 4.   Circles Task Stimuli

The N-back task is an auditory working memory task that requires participants to listen to a sequence of digits presented once every few seconds (Klatzky, et al., 2008; Reimer, 2009). Participants were required to say aloud the digit that was presented in the N-back position. An example is presented in Table 2. Participants responded after each digit presentation, except at the beginning of the stream in the more difficult conditions. In the 0-back condition, participants always repeated the digit just presented. In the 1-back condition, participants said nothing after the first digit and subsequently responded with the digit previously presented. In the 2-back condition, participants said nothing following the first two digits, and then responded with the digit presented 2 positions back in the sequence.

Table 2. Example of N-back Task Conditions

Task Condition		**Digit Sequence**
0-back	Stimulus	1 2 3 4 5 6 7 8
	Response	1 2 3 4 5 6 7 8
1-back	Stimulus	1 2 3 4 5 6 7 8
	Response	. 1 2 3 4 5 6 7
2-back	Stimulus	1 2 3 4 5 6 7 8
	Response	. . 1 2 3 4 5 6

In this experiment, N was either 1 or 2. The 2-back task was more difficult than the 1-back task because it required drivers to remember two digits at any point in time, while the 1-back required the participant to remember one digit. Because there is no visual or manual component, the interference associated with this task is primarily cognitive. This task was included in the study to determine how sensitive our measures were to the effects of cognitive distraction.

The third task category consisted of two self-paced tasks performed using the Honda Odyssey navigation system, including destination entry by street address (ADD) and destination entry by selecting a previous destination (PD). The ADD task required participants to enter addresses, which were presented one at a time on a stimulus touch screen located inside the test vehicle to the right of the navigation system (see Figure 5). For each destination, the participant performed the following sequence of operations: (1) select Enter destination by Address button, (2) press Street button, (3) enter letters of street name via touch-screen keyboard until a list was automatically generated, (4) select the street name from a list of streets, and (5) enter the street number via keyboard. After each address was entered, the participant touched the stimulus touch screen. This recorded the time to complete the address entry and displayed the next destination. The PD task required participants to select destinations that had previously been entered into the Honda Odyssey navigation system. Drivers performed this task repeatedly during each drive, obtaining new destinations via the stimulus touch screen. They used the following sequence of operations: (1) select Enter Destination by Previous Destination button, (2) press arrows to scroll through list, (3) select destination from list. The ADD task was more difficult than the PD task because address entry requires keyboard use while selecting previous destinations requires scrolling through a list (Ranney, Baldwin, Vasko, & Mazzae, 2009). Both tasks were performed with the navigation system's visual-manual interface.

2.3.4 Hypotheses

Based on the foregoing, we hypothesized that tasks performed with a visual-manual interface (Circles and navigation system tasks), which require physical manipulation of controls and visual examination of displays, will negatively affect measures of vehicle control, including lateral control and steering entropy more than tasks performed with hands-free auditory/vocal interfaces.

Jamson and Merat (2005) found improvements in lateral performance (i.e., reduced lane position variability) with cognitive secondary tasks, relative to visual-manual tasks. They hypothesized that this "improvement" was an incidental byproduct of the increase in gaze concentration to the road center that occurs with an increase in cognitive load, which was demonstrated by Victor, Harbluk and Engstrom (2005). According to this explanation, the ‘cognitive narrowing’ results in increased focus of visual resources on the road center. The resulting increase in perception of the roadway allowed an improvement in lane keeping performance. This improvement was possible because among experienced drivers lane keeping requires very little conscious attention and can be done based on peripheral inputs. Jamson and Merat (2005) hypothesized further that if this model is accurate, there should be a cost associated with the concentrated attention with increased cognitive load, namely a reduction in peripheral object detection. They were unable to test this hypothesis because their experiment did not include a peripheral detection task; however the ability to test this hypothesis is one reason to include a peripheral component to the detection task. Accordingly, we included the MDT, which uses 6 target locations, at three different eccentricities to test this hypothesis. The HDT eliminates variability among performance metrics due to the changing position of the driver’s head, which alters the target detection task when targets are presented at fixed locations in the driving scene. Accordingly, it was hypothesized that the HDT would be more sensitive to cognitive distraction effects than the MDT, which like the traditional PDT, does not control for the changes in the driver’s head position.

2.4 Method

2.4.1 Participants

Forty drivers (age range 25 to 50, mean 37.9 years old) participated in Experiment 1. Participants were recruited through advertisements placed in local newspapers and screened to ensure that they were active drivers with a valid driver’s license and a minimum of 7,000 miles driven per year. All participants reported having experience using a wireless phone while driving. Wireless phone use was considered to be a surrogate for multi-tasking; we expected drivers who were experienced phone users to be more representative of drivers who would chose to perform various secondary tasks while driving. Fifty-seven percent of the participants reported some previous experience with a navigation system. Data for Experiment 1 were collected between March and May of 2009.

2.4.2 Laboratory

Experiment 1 was conducted in 2009 in the TRC Data Collection Annex, located in a light industrial/commercial development in Plain City, Ohio. The leased space consisted of a 25 ft x 40 ft commercial garage with a high ceiling and no windows. The garage was connected by a hallway to a pair of offices, a restroom and the participant entrance. The front office was used to interview participants.

2.4.3 Apparatus

Components of the fixed-base simulator included a production test vehicle (2007 Honda Odyssey Touring), an Intel Pentium 4 computer, a ceiling-mounted digital projector (1024 x 768) positioned on top of the vehicle, and a forward projection screen (10 ft x 8 ft), which was located approximately 12 feet in front of the driver’s seated position. A touch screen was installed inside the vehicle and was connected to a separate computer, which was used to generate stimuli for secondary tasks (see Figure 5).

Figure 5. Odyssey Interior and Touch Screen Showing Secondary Task Stimulus

Sensors that recorded steering, accelerator and brake inputs were attached temporarily to the test vehicle. Specifically, a bracket (see Figure 6) was developed to couple the front tire of the test vehicle to a turn plate on the ground while the vehicle tires were off the ground (vehicle supported by 5 jack stands) (see Figure 7). This allowed drivers to sit inside the Odyssey while operating the driving simulator. The bracket and turn plate assembly mounted to the front tire 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.

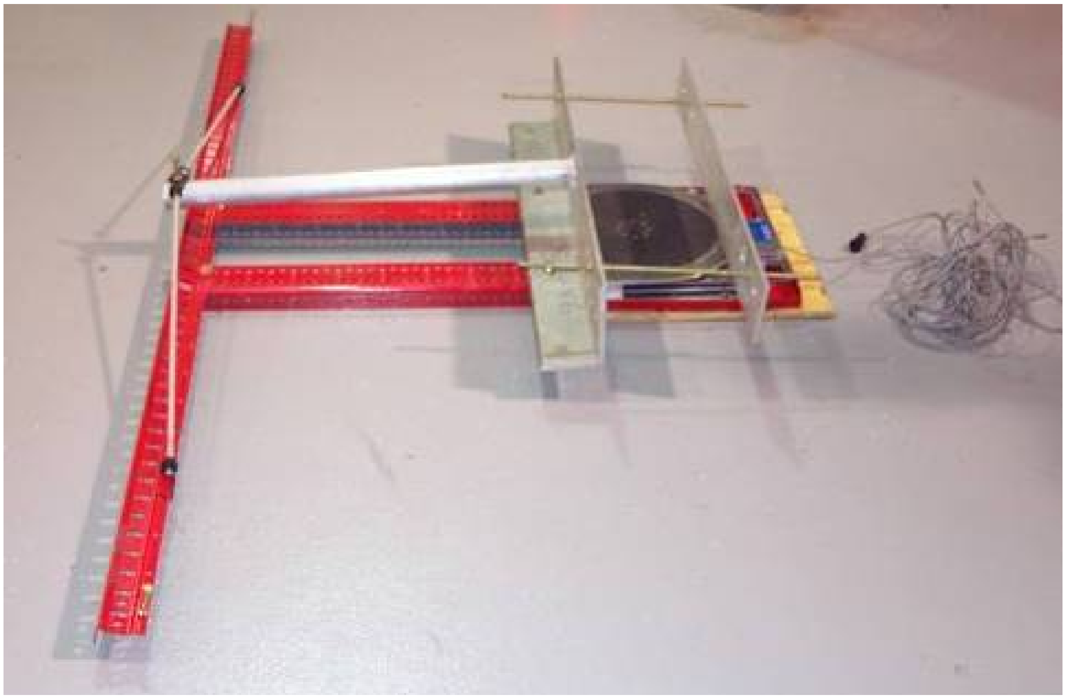


Figure 6. Apparatus for Recording Steering Wheel Movement

Figure 7. Steering Apparatus Installed on Front Tire

The Subject Vehicle (SV) MicroDAS data acquisition system (Barickman & Goodman, 1999) for Experiment 1 was configured to collect hand wheel position, brake and throttle inputs, and participant responses to the MDT and HDT. In addition, the STISIM simulation computer collected data for its respective performance measures. The primary data channels for Experiment 1 are displayed in Table 3.

Table 3. Subject Vehicle Data Collection Channels for Experiments 1 and 3

Data Channel	Description	Units	Resolution
Vehicle Speed	STISIM	km/h	1 km/h
Range	Distance to the LV, STISIM	m	.5 m
Range-Rate	Relative velocity between the SV and the LV, STISIM	m/s	.1 m/s
Lateral Position	Lateral position of the SV in reference to the simulated lanes, STISIM	cm	2 cm
Hand Wheel Position	Angular position of the steering wheel (0 degrees = straight)	deg	.1 deg
UTC Time	Time of day	HH:MM:SS	1 s
Event Task	DT button press response	0 or 1	1/30th s

The simulator plus secondary task setup is shown in Figure 8.

Figure 8. Simulator with Secondary Task

A Seeing Machines FaceLAB eye tracking system was used to record the driver's head pose and gaze. Head pose uses three parameters to define position and three parameters to define orientation. FaceLAB outputs gaze rays for each eye. Each ray has an origin at the center of the respective eye and vectors pointing toward the object being looked at. Gaze is represented as pitch and yaw angles. The pitch and yaw angles are transformed into a direction vector. Dual gaze is converted into a single gaze vector. The system used two stereo cameras mounted on the dashboard and was relatively unobtrusive. To assist the system in tracking facial features, participants applied five latex target stickers to their faces during system calibration.

2.4.4 Procedure

Each participant completed one session, lasting approximately four hours. Upon arrival, the participant was asked to read and sign the Participant Information Summary (See Appendix C), thereby giving informed consent to participate in the study. No individuals declined to participate.

The participant was escorted to the experimental vehicle and given an overview of the vehicle controls and displays, including adjusting the seat and steering wheel. This was followed by an explanation of the monetary performance incentive system (see section 2.4.5) and the Rating Scale Mental Effort (RSME) (See Appendix D). The participant was then asked to affix the latex stickers to his or her face for eye tracker calibration. During this procedure, the experimenter instructed the participant concerning head position and point of gaze. Eye tracker calibration was completed.

Next, the participant was given instructions and practice for the driving task components, including the MDT or HDT (the between-subjects factor). The participant was then given an opportunity to ask questions about any aspect of the protocol.

Training on the secondary tasks began following a break. To simplify the process, participants were first trained on the two difficulty levels of one type of secondary task. Once trained, they completed the two practice trials and then the two main test trials associated with that particular secondary task before moving on to the next secondary task type to repeat the training process.

Experiment 1 consisted of 7 main driving trials (2 for each of 3 secondary task types plus one baseline trial), and approximately 10 practice drives. Each main trial lasted approximately three minutes. After each trial, the experimenter asked the participant to complete the RSME and provided performance feedback. The experimenter then described the next trial and secondary task. The participant was offered a break after each block of secondary task types. The experimenters were at a control station behind the vehicle during data collection. Communication with the participant was accomplished via two-way radio.

At the completion of data collection, the participant exited the vehicle and completed a simulator sickness questionnaire (Appendix E: Simulator Sickness Questionnaire) to determine if rest was required before being allowed to drive home. The experimenter paid the participant a total of two amounts: (1) Base pay for participation, and (2) Performance incentive pay. The experimenter answered any questions and returned the participant to his or her personal vehicle.

2.4.5 Monetary Incentives

Participants were given a base pay of $26 per hour, plus monetary incentives to motivate acceptable performance. Monetary rewards were awarded based on experimenter ratings as shown in Table 4. Incentive amounts were defined to establish priorities among the three task components. For example, to emphasize driving as the highest priority, the car-following task was associated with the highest monetary values.

Table 4. Experiment 1 Incentive Amounts per Trial

		Performance		
Task	**Priority**	**Good**	**Acceptable**	**Poor**
Car Following	1	$1.80	$0.90	$0.0
Secondary Task	2	$1.40	$0.70	$0.0
HDT/MDT	3	$0.80	$0.40	$0.0
Total		$4.00	$2.00	$0.0

During each session, participants in Experiment 1 completed 7 main trials. On each of the 7 trials, the participant had the opportunity to earn $4.00. Thus, for good performance, each participant could earn an additional $28.00.

2.4.6 Data Reduction

Data from the STISIM trials were reduced to compute the following driving performance measures:

Coherence (Cohere (Freq) or Cohere (CC)). Coherence is a measure of squared correlation, which reflects the degree to which the following vehicle is able to match the periodicity of the lead vehicle speed signal. Coherence is used both as a measure of car-following performance and as a test of whether the associated measure of phase shift (car-following delay) is interpretable. The calculation of coherence requires a car-following paradigm in which the lead vehicle speed changes can be represented as a combination of sine waves. A detailed discussion of the computation of coherence is presented in Ranney et al. (2007). In this study, we explored an alternate computation approach, based on cross correlation instead of frequency analysis. Accordingly, we have two measures of coherence, which are designated as Cohere (Freq) and Cohere (CC). Details of the analyses based on cross correlation are presented in Appendix B.

Phase Shift (Delay (Freq) or Delay (CC)). This measure represents the response lag in car following. Its interpretation is similar to that of discrete response time measures in that longer delay values reflect poorer performance than shorter values. When coherence is relatively high (e.g., ≥ 0.80), the driver is adequately following the lead vehicle's speed changes, which implies that the associated measures are meaningful. When coherence values are low, the estimates of phase shift (delay) are considered suspect. We therefore included phase shift values in our analysis only for trials for which coherence was greater than 0.8. Less than 5 percent of the data were eliminated due to this problem. We also explored an alternate computational approach, based on cross correlation instead of frequency analysis (see Appendix B). Thus, we present two measures of delay, designated as Delay (Freq) and Delay (CC).

Mean Headway (M Hdwy). While driving, participants were instructed to maintain a constant following distance (headway) during all trials. Our previous work (Ranney et al., 2005), as well as that of Brookhuis (Brookhuis, De Vries, & De Waard, 1991), has shown that drivers have considerable difficulty maintaining a prescribed following distance. Thus, despite instructions, some drivers increased their following distances while performing secondary tasks. This measure has been interpreted as reflecting compensation for increased demands during secondary task performance, relative to baseline driving.

Standard Deviation of Headway (Std Hdwy). Drivers attempted to maintain a consistent following distance. This measure characterizes their success in doing so.

Standard Deviation of Lane Position (SDLP). This measure reflects the variability of lateral position over the entire data collection interval. It has been widely used as a measure of driving performance and has been shown to be sensitive to impairment due to fatigue, alcohol, drugs and distraction.

Steering Entropy (Steer Entropy). Developed by Boer (Boer, 2000), steering entropy measures the error in steering angle associated with loaded conditions (secondary task present) relative to a designated baseline run. The measure is based on autocorrelation and represents the frequency and extent of high-frequency corrections following periods when the driver's visual attention is diverted from the roadway.

MDT and HDT Mean Response Time (DT MRT). Drivers responded to approximately 20 targets during each driving trial. Responses recorded between 0.2 and 2.0 seconds following the target activation were considered correct responses. Mean response time is computed for the correctly detected targets on each trial.

MDT and HDT Proportion Correct (DT P Corr). This measure represents the proportion of DT targets detected correctly on a given trial.

Percentage of Time Viewing Road Center (PRC). Using the eye position measures provided by FaceLAB, we defined a road center area for each subject. Details of this computational procedure are presented in Appendix A (Appendix A: Analysis of FaceLAB Data). We used this area to classify all samples of eye position obtained during each trial that reached our quality criterion. The result was a measure of the proportion of driving time that the participant was focused on the road ahead for each trial.

Proportion of Long Glances (P Long Glance). The duration of each glance away from the road center was computed for each trial. We computed the proportion of glances away from center that exceeded 1.5 seconds.

2.4.7 Other Measures

We used the Rating Scale Mental Effort (RSME) workload rating scale (see Appendix D) to measure the participants' ratings of the subjective difficulty associated with each combination of primary and secondary task.

2.5 Results

We used Proc Mixed of SAS (Version 9.1.3) to compute an analysis of variance (ANOVA) for each dependent measure. The statistical model included the two between-subject factors (Detection Task and Driving Task Difficulty), each with two levels, and Secondary Task, which had the following seven levels:

1. Baseline – No secondary task
2. Circles 115/150 – Easy visual discrimination
3. Circles 130/150 – Difficult visual discrimination
4. Auditory 1-back
5. Auditory 2-back
6. Navigation system destination entry by address (ADD)
7. Navigation system select previous destination (PD)

Each secondary task had two levels, which differed in difficulty. The initial focus of the analysis was to determine which metrics had sufficient sensitivity to detect differences between the respective conditions for each secondary task. We therefore identified the following planned comparisons:

1. Auditory (cognitive): 1-back vs. 2-back
2. Circles: 115/150 vs. 130/150
3. Navigation: Manual destination entry vs. Select previous destination

Separate *F* tests were computed for each planned comparison for each performance measure. Probability values were adjusted for familywise error by using Hochberg's step-up method (Westfall, Tobias, Rom, Wolfinger, & Hochberg, 2003). Adjusted *p* values of less than .05 were considered to be statistically significant. Adjusted *p* values between .05 and .10 were considered marginal and discussed where applicable. A summary of the results of the planned comparisons with adjusted *p* values is presented in Table 5 and Table 6.

Table 5. Summary of Planned Comparisons Results from Experiment 1 (Driving Performance Measures)

Task	Comparison	Delay (Freq)	Delay (CC)	Cohere (Freq)	Cohere (CC)	M Hdwy	Std Hdwy	SDLP	Steer Entropy
N-back	1 vs. 2-back	.0081*	.0046*	(.48)	.0042*	(.11)	.0007*	(.37)	(.95)
Circle	115 vs. 130	(.64)	(.55)	(.48)	(.60)	(.75)	(.35)	(.64)	(.95)
Navigation	ADD vs. PD	.0124*	.0272*	(.28)	.0042*	(.11)	(.22)	.0026*	(.14)

* Statistically significant difference ($p < .05$)
+ Marginally significant ($.05 < p < .10$)
Parentheses denote differences that were not statistically significant

Table 6. Summary of Planned Comparisons from Experiment 1 (Visual Performance & Subjective Workload Measures)

Task	Comparison	DT MRT	DT P Corr	PRC	P Long Glance	RSME
N-back	1 vs. 2-back	.0023*	.0067*	(.66)	(.71)	< .0001*
Circle	115 vs. 130	(.51)	(.89)	(.76)	(.40)	(.11)
Navigation	ADD vs. PD	.0070*	(.56)	(.76)	(.71)	< .0001*

* Statistically significant difference ($p < .05$)
+ Marginally significant ($.05 < p < .10$)
Parentheses denote differences that were not statistically significant

A second set of analyses was done to assess metric sensitivity for detecting differences between secondary tasks. For these analyses, data from the respective task conditions of the same task were combined to create a single mean for each task. The results of the planned comparisons with adjusted *p* values are presented in Table 7 and Table 8.

Table 7. Summary of Task Group Comparisons Results from Experiment 1 (Driving Performance Measures)

Comparison	Delay (Freq)	Delay (CC)	Cohere (Freq)	Cohere (CC)	M Hdwy	Std Hdwy	SDLP	Steer Entropy
Baseline vs. N-back	<.0001*	<.0001*	<.0001*	<.0001*	.0072*	<.0001*	(.63)	.0649+
Baseline vs. Circles	.0003*	<.0001*	<.0001*	<.0001*	.0072*	<.0001*	<.0001*	<.0001*
Baseline vs. Navigation	<.0001*	<.0001*	<.0001*	<.0001*	<.0001*	<.0001*	<.0001*	<.0001*
N-back vs. Circles	(.46)	(.50)	(.68)	(.38)	(.87)	(.12)	<.0001*	<.0001*
N-back vs. Navigation	<.0001*	<.0001*	<.0001*	<.0001*	.0009*	<.0001*	<.0001*	<.0001*
Circles vs. Navigation	<.0001*	<.0001*	<.0001*	<.0001*	.0014*	<.0001*	.0002*	.0055*

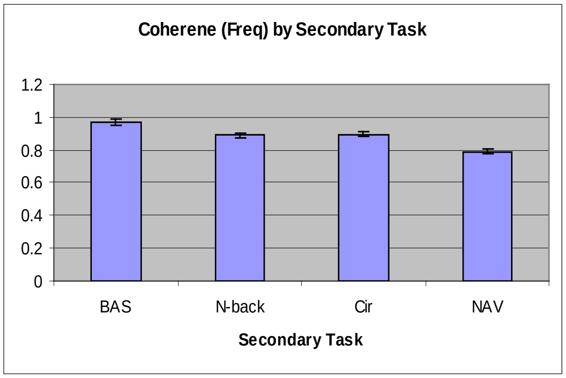
* Statistically significant difference ($p < .05$)
+ Marginally significant ($.05 < p < .10$)
Parentheses denote differences that were not statistically significant

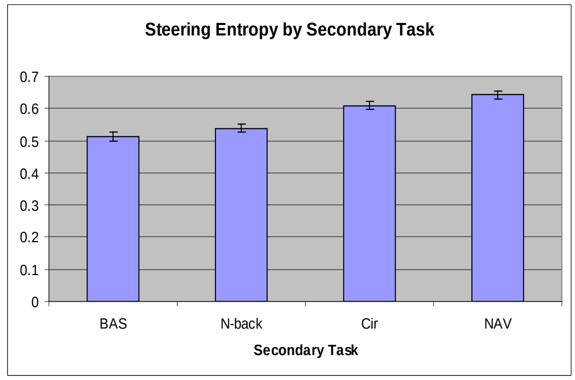
Table 8. Summary of Task Group Comparisons (Visual Performance & Subjective Workload Measures)

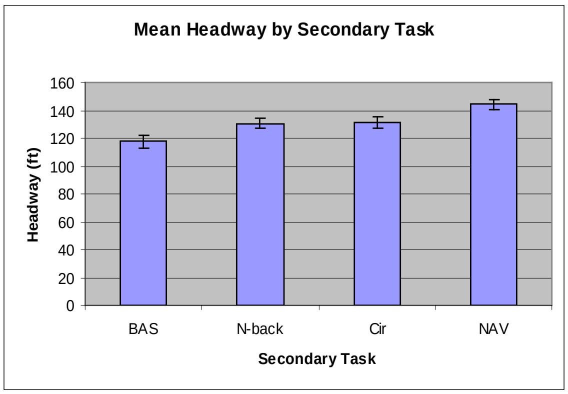
Comparison	DT MRT	DT P Corr	PRC	P Long Glance	RSME
Baseline vs. N-back	<.0001*	.0086*	.0042*	(.56)	<.0001*
Baseline vs. Circles	<.0001*	(.11)	<.0001*	<.0001*	<.0001*
Baseline vs. Navigation	<.0001*	<.0001*	<.0001*	<.0001*	<.0001*
N-back vs. Circles	(.52)	(.11)	<.0001*	<.0001*	<.0001*
N-back vs. Navigation	<.0001*	<.0001*	<.0001*	<.0001*	(.25)
Circles vs. Navigation	<.0001*	<.0001*	<.0001*	.0134*	<.0001*

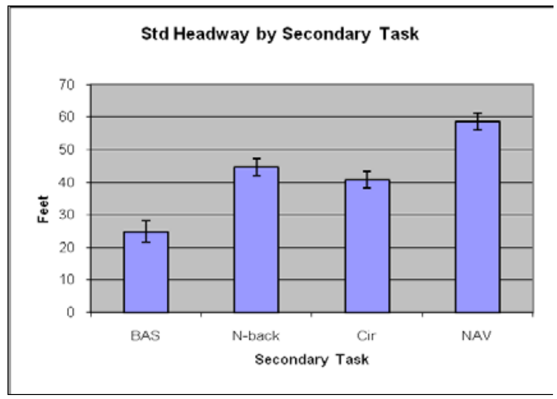
* Statistically significant difference ($p < .05$)
+ Marginally significant ($.05 < p < .10$)
Parentheses denote differences that were not statistically significant

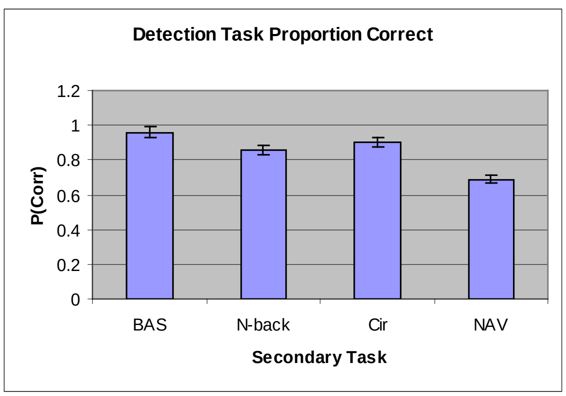
Metric means for secondary task categories are presented in Figure 9.

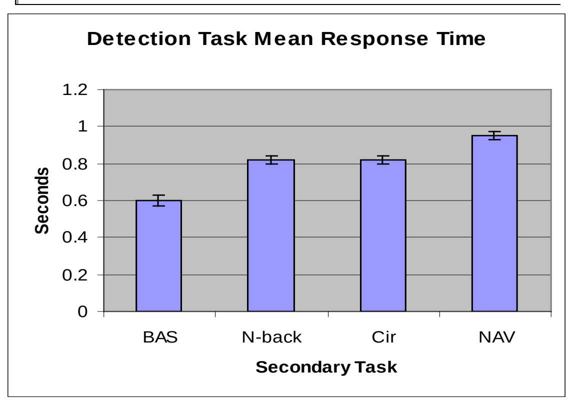

Coherene (Freq) by Secondary Task
1.2
1
0.8
0.6
0.4
0.2
0
BAS
N-back
Cir
NAV
Secondary Task

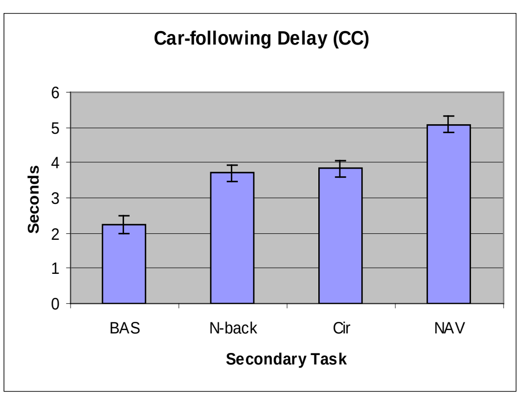

Steering Entropy by Secondary Task
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
BAS
N-back
Cir
NAV
Secondary Task

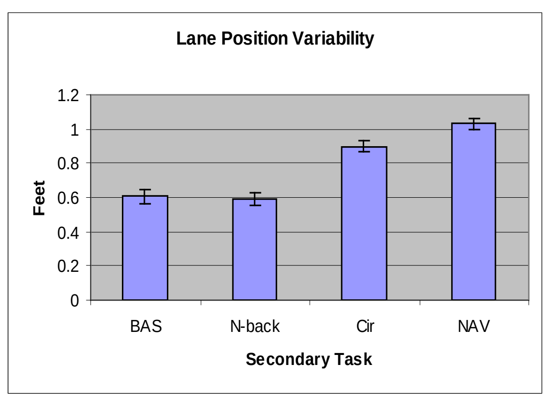

Mean Headway by Secondary Task
Headway (ft)
160
140
120
100
80
60
40
20
0
BAS
N-back
Cir
NAV
Secondary Task

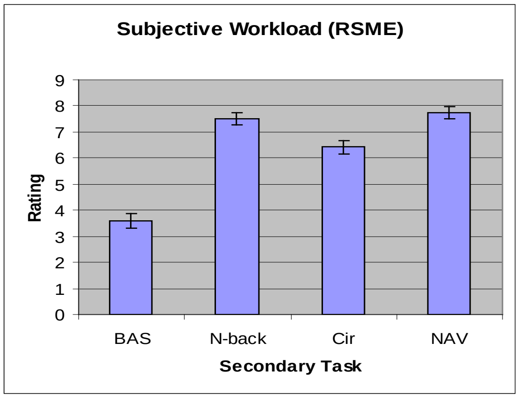

Std Headway by Secondary Task
Feet
70
60
50
40
30
20
10
0
BAS
N-back
Cir
NAV
Secondary Task


Detection Task Proportion Correct
P(Corr)
1.2
1
0.8
0.6
0.4
0.2
0
BAS
N-back
Cir
NAV
Secondary Task


Detection Task Mean Response Time
Seconds
1.2
1
0.8
0.6
0.4
0.2
0
BAS
N-back
Cir
NAV
Secondary Task


Car-following Delay (CC)
Seconds
6
5
4
3
2
1
0
BAS
N-back
Cir
NAV
Secondary Task


Lane Position Variability
Feet
1.2
1
0.8
0.6
0.4
0.2
0
BAS
N-back
Cir
NAV
Secondary Task



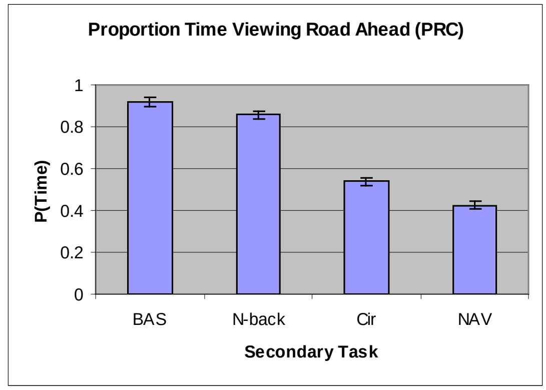


Figure 9. Between Secondary Task Category Means (± Standard error) – Experiment 1

2.6 Sensitivity of Metrics to Differences between Task Conditions

The results of the planned comparisons for within-task differences revealed significant sensitivity for the differences in the N-back (cognitive) task and the Navigation tasks, but no sensitivity for the differences in the Circles (visual-manual) task. Of particular interest with respect to the question of metric sensitivity for distraction effects that are primarily cognitive, was the finding that seven metrics were sensitive to differences between conditions in the N-back task. These metrics included both measures of detection task performance (DT MRT and DT P Corr), four measures of car-following performance (Delay (Freq), Delay (CC), Cohere (CC), and Std Hdwy), plus RSME, the subjective workload ratings. The M Hdway, SDLP, and Steer Entropy metrics were not sensitive to these differences. SDLP, which characterizes lateral position control, was, however, sensitive to the differences between Navigation tasks.

None of the metrics was sensitive to the differences between the Circles task conditions. This was true also for RSME, which indicated no differences in subjective workload. This pattern of results clearly reflects the fact that the conditions selected for the Circles task were too similar for any of the metrics. The differences between conditions in the present study (Δ= 15) were less than one third those used in a previous study (Δ= 50), in which we found consistent differences between conditions for most metrics. The limits of the metrics' sensitivity lie somewhere between these two values.

2.7 Sensitivity of Metrics to Differences between Secondary Tasks

The second set of analyses was intended to examine the sensitivity of the metrics for differences between task categories. Thus the question was whether the metrics could detect differences between Baseline measures and those associated with the different secondary tasks. As shown in Table 7 and Table 8, most metrics were sensitive to differences between the different task categories. The exception appeared to be the finding that the N-back task was in the aggregate not different from the Circles Task among measures of car-following performance and target detection. Table 9 summarizes the differences between the N-back and Circles tasks for the metrics that exhibited significant differences.

Table 9. Mean Values for Metrics with Significant Differences between N-back and Circles Tasks

Metric	N-back	Circles	Interpretation
SDLP	.59	.90	Circles more disruptive to lateral control
Steer Entropy	.54	.61	Circles more steering error
PRC	.86	.54	Circles more Eyes off road time (EORT)
P Long Glance	1.82	11.34	Circles more long glances away
RSME	7.50	6.40	N-back more demanding subjectively

These results indicate that the Circles task degraded lateral control and steering performance more than the N-back. This difference was undoubtedly associated with the decrease in time spent looking straight ahead while performing the Circles task, including a large difference in the average number of long glances away from the forward view. Despite these performance differences, the participants rated the N-back task as demanding more mental effort, overall, as reflected in the RSME differences.

2.8 Methodological Factors Hypothesized to Affect Metric Sensitivity

The experimental design included two methodological factors, detection task and car-following speed profile, which were varied between subjects. Two detection tasks were used; half of the participants had a head-mounted detection task (HDT) while the other half had a computer generated on-screen detection task involving multiple targets (MDT). Analyses focused on the effects of these factors included the two between-subject factors in the statistical model. Of primary interest were the main effects of these factors and their interactions with Secondary Task. Effects of detection task are considered first.

Analyses of all metrics included main effects of detection task plus the Secondary Task x Detection Task interaction effect. This interaction effect is of particular interest because its significance could denote an improvement in metric sensitivity associated with one of the detection tasks. Two metrics (DT MRT and Long Glance Frequency) were had significant interactions between Secondary task condition and Detection task. These interactions are explored in detail.

Detection Task Mean Response Time (DT MRT). HDT response times were generally faster than the corresponding MDT response times. This was reflected in a significant main effect of Detection Task, $F\ (1,37.5) = 11.12$, $p = .0019$. The interaction effect between Detection Task and Secondary Task Conditions was also significant, $F\ (6,216) = 2.92$, $p = .0092$. Post hoc comparisons were performed to determine whether one of the detection tasks was more sensitive to the hypothesized differences between task conditions. These comparisons are shown in the following table (Table 10), which is separated into three paired comparisons between the two detection tasks.

Table 10. Post Hoc Comparisons of Detection Task by Secondary Task Interaction Effect

Secondary Task	**Comparison**	**Detection Task**	**Adjusted P > |t|**
Auditory/Cognitive	1-Back vs. 2-Back	HDT	0.2484
Auditory/Cognitive	1-Back vs. 2-Back	MDT	0.0190*
Circles (Visual-Manual)	115/150 vs. 130/150	HDT	0.7441
Circles (Visual-Manual)	115/150 vs. 130/150	MDT	0.8439
Navigation Tasks	ADD vs. PD	HDT	0.8439
Navigation Tasks	ADD vs. PD	MDT	0.0070*

*Statistically significant difference (p < .05)

The means for each of the three comparisons are presented in the Figure 10. The results show that the two detection tasks differed in their sensitivity to detect differences in two of the three secondary task comparisons. The metric, DT MRT, was sensitive to differences between the levels of the auditory/cognitive task and the navigation tasks when the MDT was used but not when the HDT version was used.

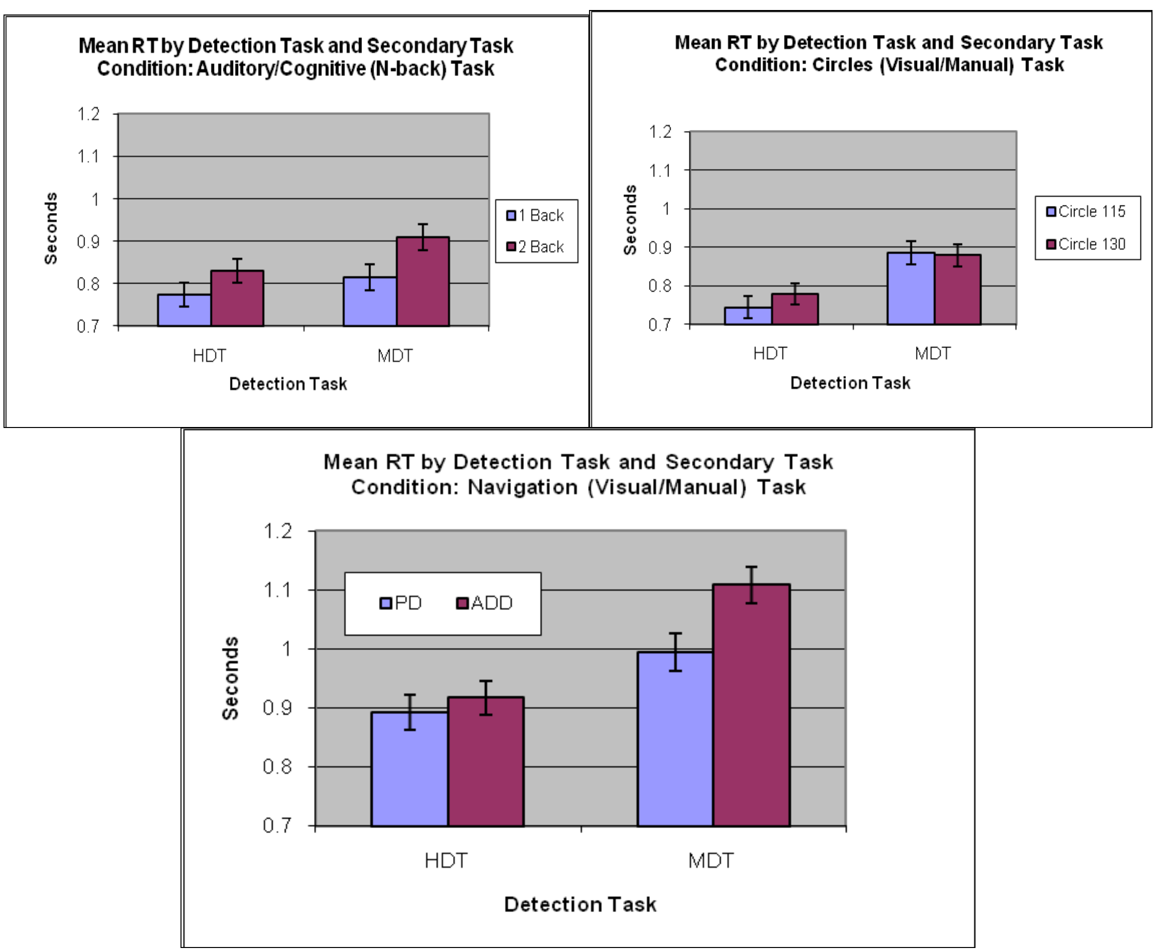


Figure 10. Detection Task Comparisons (± Standard error) – Experiment 1

Long Glance Frequency. The Detection Task x Secondary Task interaction was statistically significant for this metric, F (6,216) = 2.92, p = .0092. Examination of the specific comparisons delineated above, however, revealed no differences, which led to the conclusion that this interaction was not interpretable in the context of this question. Thus, there is no evidence suggesting that one detection task is more sensitive than the other with respect to the frequency of long glances. This is an indication that the head-mounted detection task did not influence glance behavior.

Driving task difficulty was the second between-subject factor. The experimental design included two speed profiles in the car-following task, referred to as Moderate and Difficult. Analysis results were generally consistent in showing no significant differences between the two difficulty conditions. The Speed Profile x Secondary Task interactions were examined to determine whether increasing the car-following task difficulty affected the metrics' sensitivity for detecting the hypothesized differences between secondary task conditions. One metric (Cohere (Freq)) revealed a significant interaction between the speed profile (car-following task difficulty) and secondary task condition. The interpretation is presented below.

Cohere (Freq). The Speed Profile (car-following task difficulty) x Secondary Task condition interaction was significant for this metric, F (6,223) = 2.19, p = .0445. Examination of the post hoc comparisons revealed no interpretable differences (see Table 11), which led to the conclusion that this interaction did not support the hypothesis that increasing the car-following task difficulty increases metric sensitivity.

Table 11. Post Hoc Comparisons of Speed Profile by Secondary Task Interaction Effect

Secondary Task	**Comparison**	**Speed Profile**	**Adjusted P > |t|**
Auditory/Cognitive	1-Back vs. 2-Back	Moderate	0.53
Auditory/Cognitive	1-Back vs. 2-Back	Difficult	0.53
Circles (Visual-Manual)	115/150 vs. 130/150	Moderate	0.16
Circles (Visual-Manual)	115/150 vs. 130/150	Difficult	0.41
Navigation Tasks	ADD vs. PD	Moderate	0.41
Navigation Tasks	ADD vs. PD	Difficult	0.41

3.0 EXPERIMENT 2

3.1 Background

Metrics derived from the simulator-based test used in Experiment 1 represent estimates of the expected level of driving performance degradation associated with secondary tasks at any point during their performance. These estimates can be combined with task duration estimates to characterize the total exposure to risk associated with a given task. Occlusion is a simple technique to estimate the total amount of eyes-off-road-time (EORT), which is used to estimate task completion time. The traditional occlusion technique involves periodic interruption (via visual occlusion) of the performance of a secondary task (e.g., navigation system destination entry) (Stevens, Bygrave, Brook-Carter, & Luke, 2004). The interruption is intended to simulate the real-world requirement of switching vision and attention between the secondary task and driving. Occlusion is accomplished with computer-controlled PLATO (Portable Liquid crystal Apparatus for Tachistoscopic Occlusion) goggles, which are a spectacle-mounted shuttering device with portable liquid-crystal apparatus. The lenses in these spectacles can be rapidly (e.g., 1-5 ms) and independently switched from a light-scattering, occluding state to a transparent state, in which up to 90 percent of incident light is transmitted.

The Enhanced Occlusion Technique (EOT) combines the traditional occlusion technique with a computer-generated auditory tracking task (Schindhelm & Gelau, 2008). The EOT was developed to address a methodological problem with the traditional occlusion technique. Specifically, with no processing load during the occluded intervals, the traditional technique allows participants to continue working on the secondary task, a phenomenon called blind operation (Gelau & Schindhelm, 2009). As a result, the occlusion technique does not provide a valid simulation of the disruption of IVIS task performance caused by the demands of driving. More specifically, the measures of time required to complete an IVIS task cannot be accurately estimated from the sum of the presentation intervals during occlusion trials. Evidence of this problem exists in low values of the resumability ratio (R), a measure of the ease of resumption of task performance following interruption, (Pettitt , Burnett, Bayer, & Stevens, 2006), which is also referred to as task interruptability (Noy, Lemoine, Klachan, & Burns, 2004). R is defined as the total shutter open time (TSOT) obtained during occlusion trials divided by the time required to complete the same task in a static, uninterrupted condition. Larger values of R are indicative of tasks that are more difficult to perform under conditions of task interruption, as in driving. In theory, if participants were not able to work on the IVIS task during the occluded intervals, the TSOT values derived from occlusion trials would always be greater than the static performance time, resulting in R values always greater than or equal to one. In practice, most R values are less than one, which indicates that participants are able to work constructively on the IVIS task during the occluded intervals. While some have argued that driving offers similar opportunities to continue working on an IVIS task while looking at the roadway, the absence of any load in the occlusion paradigm offers greater opportunity for this continued activity than on-road driving.

The addition of the auditory tracking task is intended to address this concern. Auditory tracking provides the same type of processing load as required to steer a vehicle, but without the visual demands that could conflict with the essential mechanism of the occlusion protocol, whereby attention to driving is simulated by restricting participants' view of the IVIS task. Gelau and Schindhelm (2009) showed that measures obtained using the EOT were more sensitive to differences between IVIS tasks than measures obtained using the traditional occlusion technique.

The main performance measures associated with this paradigm are the TSOT, which is an estimate of the total time required to complete the IVIS task under conditions of occlusion and R, which is a measure of task resumability.

The experiment used real-world IVIS tasks. Based on our previous work, we included two tasks with different levels of distraction potential. Specifically, we used a (more demanding) destination entry by address (ADD) task and a (less demanding) destination entry by selecting a previous destination (PD) task. In our previous work, we concluded that the differences in distraction potential were associated primarily with whether or not the task required use of the on-screen keyboard. The ADD task required keyboard entry while the PD task did not. We added a third task, List Search (LS), which required participants to find a designated city within a long list generated by the navigation system. This latter task was included to eliminate any potential confounds in the PD task, due to the requirement to search repeatedly (on successive trials) through the same list. The LS task created different lists of comparable length for each trial.

3.2 Experimental Design

Experiment 2 used a repeated-measures within-subjects design in which all participants received all treatment conditions. The main design factors (independent variables) included occlusion task condition (static, occlusion, EOT) and secondary task (ADD, PD, LS). In the static condition, participants completed all secondary tasks without interruption. In the occluded condition, drivers performed the same tasks subject to periodic interruption from the occlusion goggles. In the EOT condition, the occlusion goggles were used together with the auditory tracking task.

3.2.1 Hypotheses

One hypothesis evaluated in this experiment was that TSOT would be significantly increased by the addition of the auditory tracking task in the EOT condition. We also hypothesized that the task resumability metric (R) would increase in the EOT condition, relative to the traditional occlusion condition. Specifically, we hypothesized that R_{EOT} would be significantly greater than R_{OCC} and that R_{EOT} would be more sensitive to differences in demands between IVIS tasks than R_{OCC}. Finally, we predicted that subjective workload ratings would be greater for the ADD task relative to the two other IVIS tasks.

3.3 Method

3.3.1 Participants

Twenty-seven drivers (age range 26 to 51, mean 39.9 years old) participated in Experiment 2. To accommodate an aggressive schedule, we recruited from the sample of participants that had participated in Experiment 1. Because the demands of occlusion and auditory tracking were so different from those involved in driving the simulator, and because the within-subjects design was intended to minimize the influence of between-subject variability, we concluded that the participants' previous participation would have little effect on the results of Experiment 2. All participants were active drivers with a valid driver's license and a minimum of 7,000 miles driven per year. All participants had previous experience using a cell phone while driving. Wireless phone use was considered to be a surrogate for multi-tasking aptitude/propensity; we expected drivers who were experienced phone users to be more representative of drivers who would chose to perform various secondary tasks while driving. Data for Experiment 2 were collected between July and August of 2009.

3.3.2 Laboratory

Experiment 2 was conducted using the same laboratory space as in Experiment 1, specifically the TRC Data Collection Annex in Plain City, Ohio. The laboratory space was described in Section 2.4.2.

3.3.3 Apparatus

Occlusion is accomplished with computer-controlled PLATO (Portable Liquid crystal Apparatus for Tachistoscopic Occlusion) goggles, which are a spectacle-mounted shuttering device with portable liquid-crystal apparatus (see Figure 11). The lenses in these spectacles can be rapidly (e.g., 1-5 ms) and independently switched from a light-scattering, occluding state to a transparent state, in which up to 90 percent of incident light is transmitted. Timing of the change between occluded and transparent states was computer controlled. For this experiment, both intervals were set to 1.5 seconds (ISO, 2007).

Figure 11. PLATO Occlusion Goggles

IVIS tasks were performed using the OEM navigation system in a 2007 Cadillac Escalade. The navigation system had a touch screen interface. All tasks required manual inputs. Instructions were presented to participants on an LCD touch screen located to the right of the navigation system. A single control program, running on a computer located outside the vehicle, was used to control the occlusion timing, auditory tracking, stimulus presentation and recording of participant timing inputs from which total task times were computed.

The auditory tracking task was designed to replicate the task used by Gelau and Schindhelm (2009). Participants moved a joystick in response to tones presented in one of two audio channels (left or right). The sound location provided feedback about the participant's position on a hypothetical winding path, which was not visible to the participant. Specifically, a tone in the left channel indicated that the tracking cursor (hypothetical vehicle) was off the path to the left; a right-channel tone indicated tracking error to the right. The correct response was to move the cursor in the direction opposite the tone, in effect to move the cursor back onto the path. The feedback tone frequency increased as the magnitude of error (defined as the distance between cursor and target position) increased. The auditory feedback was presented via headphones, which were determined through pilot work to provide better sound localization than speakers. Participants used a joystick to make tracking inputs. The joystick was a "Logitech Dual Action" joystick with a mild restoring force which provides some "on center" feel. The joystick was utilized as a rate controller where deflection of the joystick determined the rate of change of the signal being controlled. The experimenter's view of the control program is shown in Figure 12, which is a time-based trace. The occlusion condition is indicated by the square-wave trace at the top of the figure. The tracking target (imaginary winding path) is indicated by the complex waveform in the upper central region of the display. As the trial progresses, the topmost vertical bar moves from left to right to indicate the position of the cursor and the occlusion condition. The width of vertical bar in the lower part of the figure represents the size of the region of acceptable performance (RAP) and the central line within that bar indicates the position of the cursor relative to this region. When the cursor remained inside the RAP there was no auditory feedback.

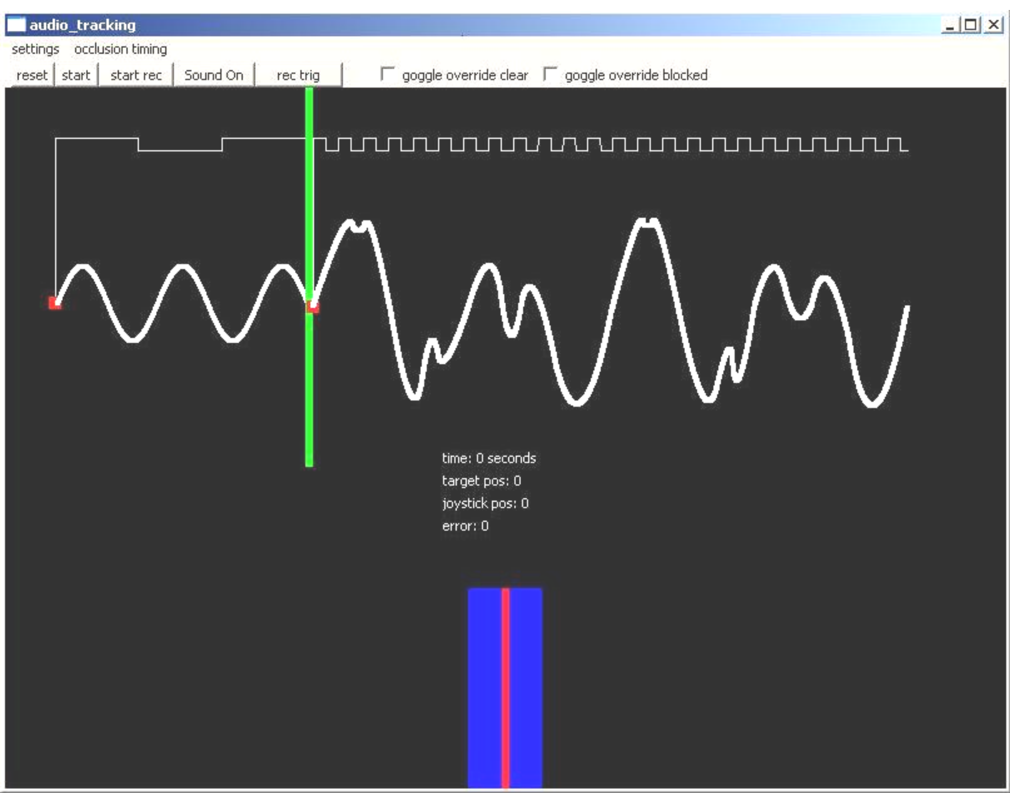


Figure 12. Occlusion and Auditory Tracking Control Program Monitor

A small video camera was positioned behind the participant so that the navigation system inputs could be recorded. This video was used to identify instances of significant system response delay (SRD).

3.3.4 Procedure

Each participant completed one session, lasting approximately four hours. Upon arrival, the participant was asked to read and sign the Participant Information Summary, thereby giving informed consent to participate in the study. No individuals declined to participate.

The participant was escorted to the experimental vehicle. When seated, the participant was given instructions for adjusting the seat so that the navigation system controls were within easy reach. This was followed by an explanation of the monetary performance incentive system (see Section 3.3.7) and the Rating Scale for Mental Effort (Appendix D: Rating Scale Mental Effort (RSME).

The experimenters monitored the experiment from a bench that was located immediately behind the experimental vehicle. Communication with the participant was accomplished via two-way radio.

At the beginning of the experiment, the Circles task was used to familiarize participants with the occlusion technique and the auditory tracking task. The Circles task was a simpler version of the task used in Experiment 1. Participants were instructed not to work on secondary tasks during the occlusion intervals.

Each participant completed 25 trials in each of the three secondary task conditions described below in Section 3.3.6. Each block of 25 trials was divided into 4 components as shown in Table 12.

Table 12. Sequence of Trials (Experiment 2)

Component	# Trials	Description
1	7	1 training, 6 practice trials, 2 in each condition (static, occlusion, EOT)
2	6	1 practice, 5 main trials in condition 1
3	6	1 practice, 5 main trials in condition 2
4	6	1 practice, 5 main trials in condition 3

Order of conditions (static, occluded, EOT) was balanced across blocks and participants so that the entire design had an equal number with each order.

Data from the 15 main trials for each of the three trial blocks were analyzed. Thus, each participant completed 75 trials of which 45 were used for analysis. After each trial, the experimenter asked the participant to complete the RSME and provided performance feedback. The experimenter then described the next trial and read secondary task instructions aloud. The participant was given a break after each block of trials.

At the completion of data collection, the experimenter paid the participant a total of two amounts: (1) Base pay for participation, and (2) Performance incentive pay. The experimenter answered any questions and returned the participant to his or her personal vehicle.

3.3.5 Occlusion Task Conditions

The experiment used three occlusion task conditions, including static, occlusion, and EOT. The static condition required participants to complete the secondary tasks without interruption. The occlusion condition required participants to complete secondary tasks while wearing occlusion goggles that alternated between opaque and transparent. The control timing was set so that both the occlusion and transparent intervals were 1.5 seconds. The EOT condition combined visual occlusion with an auditory tracking task, as described above.

3.3.6 Secondary Tasks

Secondary tasks used in this experiment were performed using the visual-manual interface of the Cadillac Navigation Infotainment System. Details of the three tasks are presented below.

Destination Entry by Address (ADD). This self-paced task required participants to use the Cadillac Navigation Infotainment System to enter an address, which consisted of a state, city, street name, and house number. Inputs were entered using a virtual, touch-screen keyboard, which required the participant to spell out the beginning letters until a list was automatically generated by the system. At this point, the participant searched the list and the selected the correct entry. Destinations were matched as closely as possible for their entry requirements. For example, addresses that prompted system users to navigate through a list using arrow keys or to provide additional information such as a street type, were eliminated. Matching was based on a task analysis, which counted screen touches. Address entry trials were organized into three groups, one for each condition (static, occlusion, EOT). Based on their rank order from initial pilot tests, the addresses were grouped so that the groups would have statistically similar means in the static condition. Individual trials were also within one standard deviation of the average total task time. The grouping was done as an additional balancing method.

List Search by City (LS). This self-paced task required participants to select a city from a larger list of cities. Participants began the task by selecting a designated state. Next, participants were instructed to input the first letter of the city and to press a “list” button to prompt the system to create and display a list of all cities within the specified state that began with the specified letter. City/state combinations were selected that created lists of approximately equal length. Participants used the scrolling feature to navigate through 5-7 subsets of the list to find the target city. The cities were presented in alphabetical order in the list.

Selecting Previous Destination (PD). This self-paced task required participants to use the Cadillac Navigation Infotainment System to select a sequence of destinations that had previously been entered into the system. Participants selected the ‘Previous Destination’ button on the touch screen, which activated a list and then scrolled through this list to find the specified destination. The Cadillac Navigation Infotainment System allowed space for 20 previously stored destinations. Because this task can potentially be completed very quickly, each trial consisted of a pair of previous destinations, which required participants to search twice through the list. Care was taken in the selection of stimuli to ensure that the location of each stimulus was such that the number of steps required to find it in the list was approximately equal.

3.3.7 Monetary Incentives

Participants were given a base pay of $26 per hour plus monetary incentives to motivate acceptable performance. Monetary rewards were awarded based on experimenter ratings as shown in Table 13. Ratings were based on the number of errors and speed of responses. Participants completed 45 trials. Thus, based on a three-hour data collection protocol, for consistently good performance, participants had the opportunity to earn approximately $9.00 per hour in addition to the base pay.

Table 13. Incentive Amounts per Trial (Experiment 2)

	Performance		
Condition	**Good**	**Acceptable**	**Poor**
Static	$0.50	$0.25	$0.00
Occlusion	$0.60	$0.30	$0.00
EOT	$0.70	$0.35	$0.00

3.3.8 Data Reduction

Data from the experiment were reduced to create the following performance measures:

Total Shutter Open Time (TSOT). TSOT values were computed for each trial in the occluded and EOT conditions by multiplying the Total Task Time (TTT) by the proportion of samples that were unoccluded.

Task Resumability Index (R). R is defined as TSOT/TTT, where TSOT is the total shutter open time defined above and TTT is the total task time recorded in the static condition. Two values of R were computed, one using TSOT values obtained from the Occlusion condition and one using TSOT values from the EOT condition. TSOT and TTT values were median values from each subject in each of the three conditions. Two R values were computed for each subject in each Secondary Task condition.

Absolute Mean Deviation (AMD). This measure of auditory tracking performance error was computed only for the EOT trials. For each sample the difference between the cursor position and the target signal position was determined. The mean was then computed using the absolute values of the deviations.

Percentage of Time Inside Region of Acceptable Performance (P RAP). As discussed above, the RAP was defined as a region of acceptable tracking performance in the EOT condition. When the cursor was inside the RAP, no auditory feedback was provided. For each sample it was determined whether the cursor was inside or outside the RAP. This measure was computed as the number of samples in which the cursor was inside the RAP divided by the total number of samples in the trial.

Rating Scale Mental Effort (RSME). Subjective workload ratings were recorded for each of the 45 trials for each participant (see Appendix D).

3.4 Results

We used Proc Mixed of SAS (Version 9.1.3) to compute an analysis of variance (ANOVA) for each dependent measure. Dependent measures included the following:

Total Shutter Open Time (TSOT). Differences were apparent across secondary task conditions, F (2,1180) = 1362.54, p < .0001 and Occlusion conditions, F (2,1180) = 92.80, p < .0001. The Secondary Task x Occlusion condition interaction was also significant, F(4,1180) = 7.40, p < .0001, reflecting the fact that unlike all other comparisons, the difference between the Static and EOT conditions was not significantly different for the Previous Destination Secondary task, t (1180) = 1.47, p = .1430. Means for this interaction effect are shown in Figure 13.

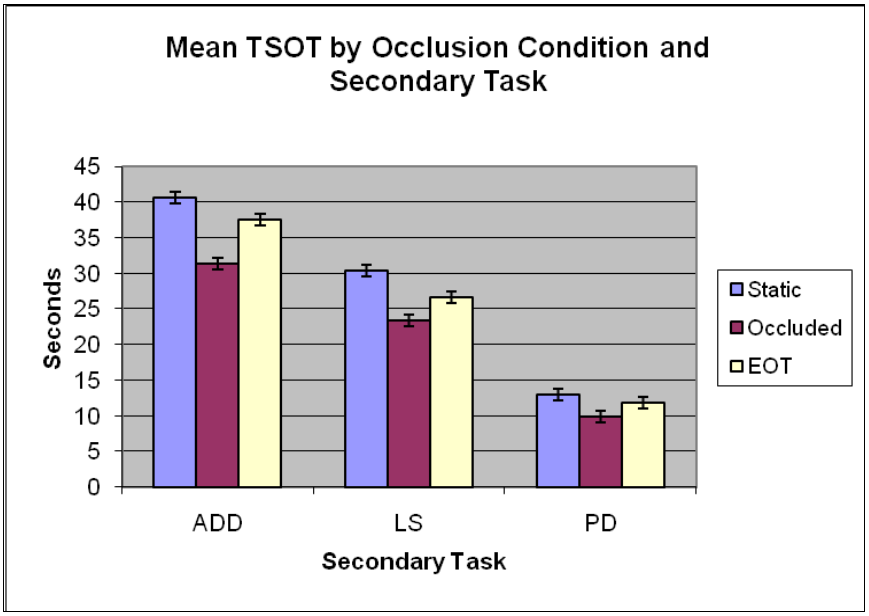


Figure 13. Total Shutter Open Time (TSOT) Means (± Standard error)

R (Task Resumability). Effects included in the ANOVA included Occlusion condition (OCC, EOT) and Secondary Task (AD, PD, LS). Results indicated significant differences between the two occlusion conditions, F (1,130) = 72.76, p < .0001 (see Figure 14). There were no differences between Secondary Task conditions, F (2,130) = 0.27, p = 0.7615, nor was the Secondary Task by Occlusion Condition interaction significant, F (2,130) = 0.72, p = 0.4908.

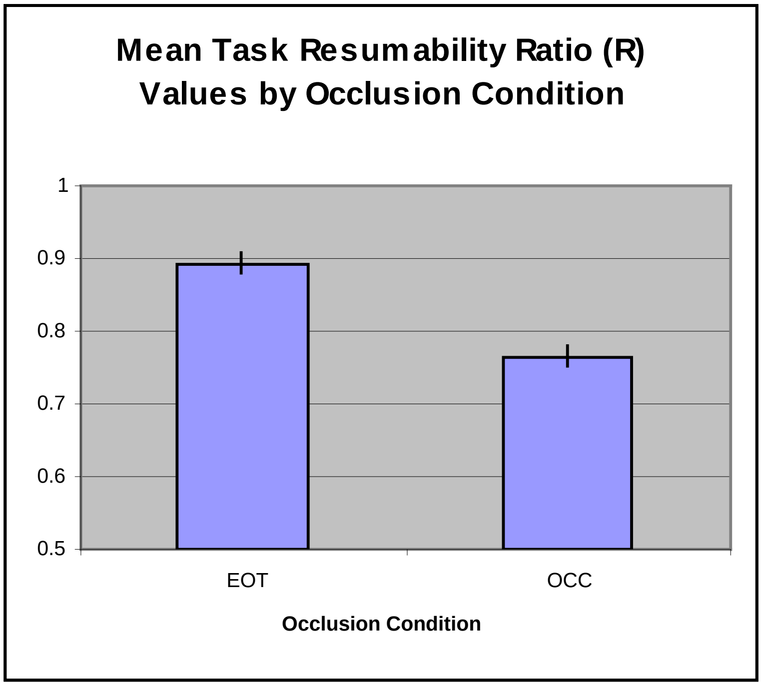


Figure 14. Means R Values by Occlusion Condition (± Standard error)

Absolute Mean Deviation. Tracking performance data was available only for the EOT trials. Thus, Secondary Task was the only factor in the ANOVA model. Results indicated a significant effect of Secondary Task, F (2,376) = 6.74, p = 0.0013. Means are presented in Figure 15. Post hoc analyses results are summarized in Table 14.

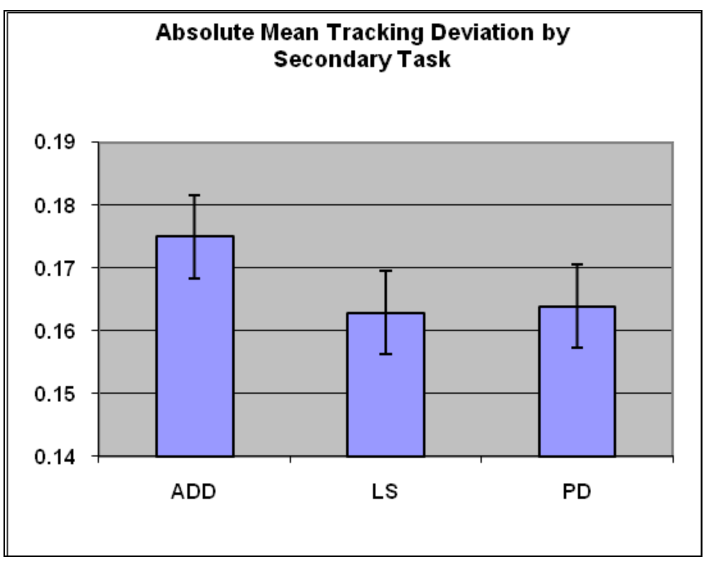


Figure 15. Absolute Mean Tracking Deviation by Secondary Task (± Standard error)

Table 14. Post Hoc Results for Absolute Mean Tracking Deviation

Comparison	Degrees of Freedom	F	p
ADD vs. LS	1,376	10.93	0.0031
ADD vs. PD	1,376	9.22	0.0051
LS vs. PD	1,376	0.07	0.7878

Percentage of Time Inside RAP. As with the previous measure, Secondary Task was the only factor in the ANOVA model. Results indicated a significant effect of Secondary Task, F (2,376) = 4.91, p = 0.0079. Means are presented in Figure 16. Post hoc analyses results are summarized in Table 15.

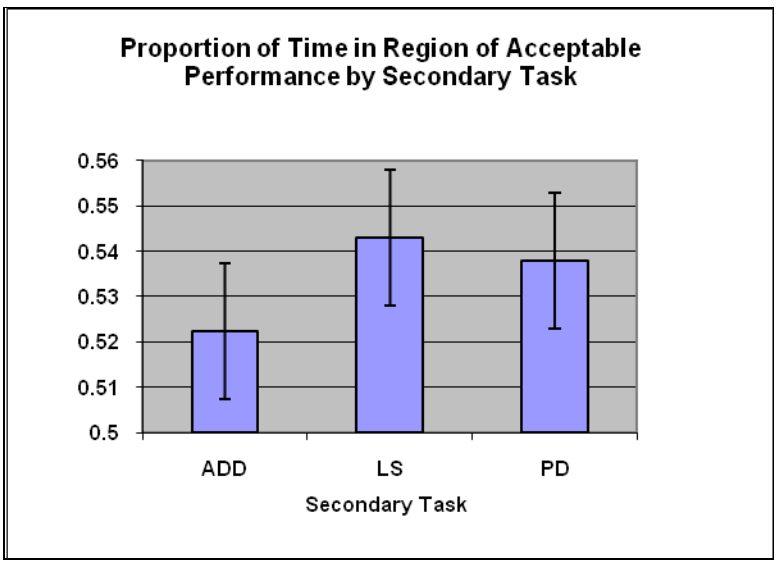


Figure 16. Means for Proportion of Time in RAP (± Standard error)

Table 15. Post Hoc Results for Percentage of Time in RAP

Comparison	Degrees of Freedom	F	P
ADD vs. LS	1,376	9.05	0.0028
ADD vs. PD	1,376	5.13	0.0241
LS vs. PD	1,376	0.55	0.4575

RSME. Analysis of RSME workload ratings included Secondary Task and Occlusion Condition as model factors. Main effects of Secondary Task, F (2,1178) = 31.29, p < 0.0001 and Occlusion Condition, F (2,1178) = 1708.87, p < .0001, were both statistically significant, as was the interaction between these two conditions, F (4,1178) = 6.40, p < .0001. The means for this latter effect are presented in Figure 17.

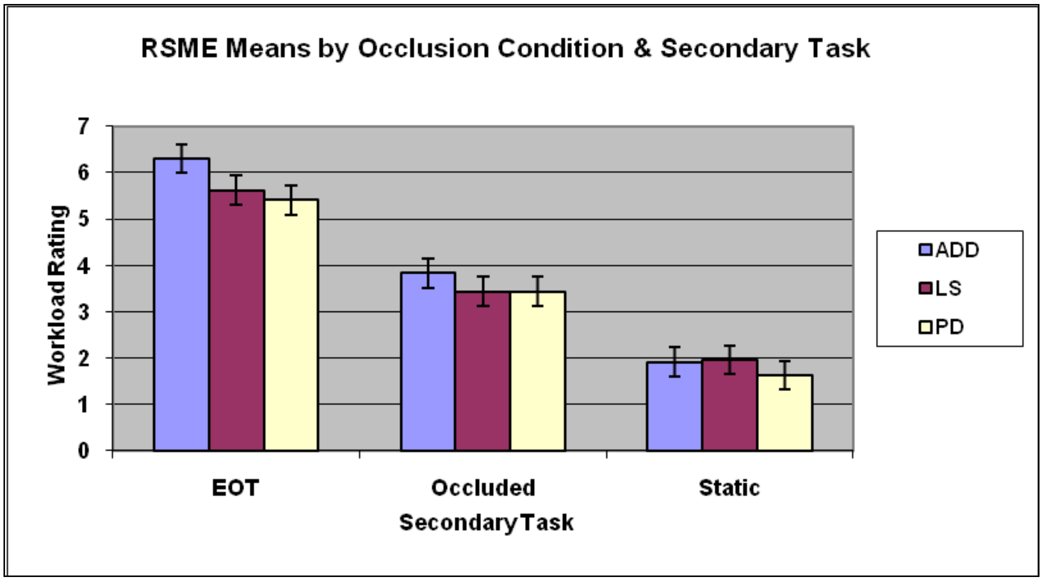


Figure 17. RSME Means by Occlusion Condition and Secondary Task (± Standard error)

Post hoc analyses were conducted to determine whether the workload ratings revealed differential effects in the various occlusion conditions. Because all paired comparisons were of interest, no subset of planned comparisons was used. The results of these analyses are presented in Table 16.

Table 16. RSME Post Hoc Results for Secondary Task x Occlusion Condition Interaction

Comparison	Degrees of Freedom	t	P
EOT: ADD vs. LS	1178	5.80	< 0.0001*
EOT: ADD vs. PD	1178	7.64	< 0.0001*
EOT: LS vs. PD	1178	1.83	0.2018
Occlusion: ADD vs. LS	1178	3.39	0.0043*
Occlusion: ADD vs. PD	1178	3.39	0.0043*
Occlusion: LS vs. PD	1178	0.00	1.000
Static: ADD vs. LS	1178	-.032	1.0000
Static: ADD vs. PD	1178	2.44	0.0594+
Static: LS vs. PD	1178	2.76	0.0297*

* Statistically significant difference ($p < .05$)
\+ Marginally significant ($.05 < p < .10$)

3.5 Discussion

TSOT values obtained in the Occlusion condition were significantly shorter than the TTT values obtained in the Static condition, reflecting the tendency of participants to continue working on the IVIS tasks during the occluded intervals. The TSOT values observed in the EOT condition were significantly greater than those from the Occlusion condition, but still somewhat less than the Static TTT values. Thus, while the auditory tracking load reduced the amount of blind operation, it did not eliminate it completely. The observation that blind operation occurs in the EOT condition suggests that blind operation is likely to exist in real-world driving. As hypothesized, the R values differed between the Occlusion and EOT conditions, due to the addition of the auditory tracking load in the latter condition. The mean for the occlusion condition (R_{OCC}) (M = 0.77) was smaller than that for the EOT condition (R_{EOT}) (M = 0.89). Because these values represent proportions of Static TTT, they suggest that participants complete approximately 23 percent (1.0 – 0.77 = 0.23) of the task during occluded intervals in the Occlusion condition and 11 percent (1.0 – 0.89 = 0.11) in the occluded intervals during the EOT condition.

Unlike Schindhelm and Gelau (2009), R_{EOT} was not more sensitive to differences between the IVIS task conditions than R_{OCC}. There were no differences between task conditions for either R value, indicating that the IVIS task conditions did not differ with respect to task resumability. Interestingly, the IVIS tasks did differ with respect to their associated tracking performance. Both measures of tracking performance, the absolute mean deviation (AMD) and the proportion of time in the Region of Acceptable Performance (P RAP), exhibited differences between the ADD and other tasks. The PD and LS task were not different for this measure. Differences between ADD and other task conditions suggest that performing the tasks did have differential effects on tracking performance. The differences were in the direction expected, indicating that the ADD task was most disruptive to tracking performance. Thus, while occlusion alone does not adequately simulate the driving task demands, the addition of a tracking task introduces the possibility that participants will devote differential amounts of attention to the tracking task for different secondary task. This finding underscores the importance of considering both tasks when analyzing performance in dual task situations.

RSME subjective workload ratings exhibited a pattern of differences suggesting that participants' assessments of workload differed between the Occlusion and Static conditions. Specifically, participants rated the ADD task to be more demanding than the other two tasks when performed under conditions of interruption (Occlusion and EOT conditions), relative to the Static condition. This finding suggests that the ADD task should exhibit higher R values than the other two tasks, since R purportedly represents the relative difficulty of performing a task under conditions of task interruption. However, this difference was not observed, which raises questions about the usefulness of the R metric and/or the quality of the data used to compute R values.

The primary contribution of the occlusion paradigm is the estimate of TSOT, which represents the total amount of time required to perform a task under conditions of task interruption. The present results provided two estimates of TSOT, one for occlusion alone and one for occlusion with auditory tracking (EOT). Both values were smaller than the TTT obtained in the Static condition, supporting the conclusion that even when performing a tracking task with a load that resembles the demands of steering, participants were able to work constructively on IVIS tasks during the occluded intervals. This raises the question of how best to estimate the time required to perform a specific task. This problem is complicated by the fact that most complex tasks, particularly those involving navigation systems, can differ considerably in their durations depending on the specifics of the destination to be entered or found in a database.

4.0 EXPERIMENT 3

The objective of Experiment 3 was to use the modified test protocol to assess the distraction potential of multiple (in-vehicle and portable) systems with comparable capabilities. Test protocol modifications were made based on the results of Experiment 1. Specifically, we incorporated the multiple target detection task (MDT), which was found in Experiment 1 to provide greater sensitivity for detecting effects of cognitive distraction. It also provided the practical benefit of not interfering with the quality of the eye position data. Because the results of Experiment 1 revealed no consistent differences between the different levels of driving task demand, we retained the level that had been used in our previous work, to allow comparison of results across studies. The revised test protocol was used to evaluate the distraction potential of IVIS tasks performed with three navigation systems, including one in-vehicle system and two portable systems. Participants performed two tasks with each system, including destination entry by address (ADD) and selecting a previous destination (PD). The ADD task has generally been found to be more disruptive to driving than the PD task both in our work and elsewhere.

Three navigation systems were used including one OEM system and two portable GPS systems. The portable systems were selected based on usability ratings in Consumer Reports (2009). Because the focus of this work is on determining whether usability ratings predict distraction potential and not on comparing specific GPS systems, the specific systems will not be identified. Rather they will be referred to as the High-Usability and Low-Usability systems. A summary of the comparative ratings in selected categories is presented in the following table:

Table 17. GPS Usability Ratings (Consumer Reports)

System	Entering Destination	Info for driver	Use of controls	Display
High-Usability	5	4	5	5
Low-Usability	3	4	3	5

Ratings were 1-5, with 5 being best. The High-Usability system had higher ratings for entering a destination and use of controls than the Low-Usability system.

4.1 Experimental Design

Experiment 3 used a within-subjects design in which all participants received all treatment conditions. Experimental factors included two levels of secondary task (ADD, PD) and three navigation systems. Each participant thus received each combination of these two factors. In addition to the secondary task trials, each participant completed an initial and final baseline trial in which there was no secondary task. Thus, excluding training and practice, each participant completed eight three-minute drives.

4.2 Method

4.2.1 Participants

Thirty-six drivers (age range 26 to 56, mean 43.9 years old) participated in Experiment 3. Participants differed from those used in Experiments 1 and 2. They were recruited through advertisements placed in local newspapers and screened to ensure that they were active drivers with a valid driver's license and a minimum of 7,000 miles driven per year. All participants reported experience using a wireless phone while driving. Wireless phone use was considered to be a surrogate for multi-tasking aptitude/propensity; we expected drivers who were experienced phone users to be more representative of drivers who would chose to perform secondary tasks like those used in the present Experiment while driving. Seventy-eight percent of the participants reported previous experience using a navigation system. Data for Experiment 3 were collected between September and October of 2009.

4.2.2 Laboratory

Experiment 3 was conducted in the TRC Data Collection Annex in Plain City, Ohio. The space was described in Section 2.4.2.

4.2.3 Apparatus

Components of the fixed-base simulator and data acquisition system are the same as were used in Experiment 1. They described in Section 2.4.3.

Subject Vehicle. The same Honda Odyssey Touring mini-van used in Experiment 1 was used in Experiment 3. The MicroDAS data acquisition system configuration was the same as was used in Experiment 1. The primary data collection channels are displayed in Table 3.

Navigation Systems. Descriptions of the navigation systems are presented in the following sections. Additional pictures of the three navigation systems used in Experiment 3 are presented in Appendix G: Display Screen Images of 3 Navigation Systems, Experiment 3

OEM Navigation System. The factory-installed navigation system consisted of a touch screen in combination with a set of physical buttons surrounding the screen (see Figure 18).

Figure 18. OEM Navigation System

When using the OEM navigation system, selections were made either by pressing a button shown on the touch screen or by toggling a joystick located to the left of the touch screen (ENT) to highlight the desired button and then pressing in on the joystick to activate the highlighted button. If an error was made, the participant was instructed to press the 'CANCEL' button, which was located above the joystick to the left of the touch screen. This activated the previous screen on the display. To return to the main menu, the participant pressed the 'MENU' button, located to the right of the touch screen.

<u>High-Usability Portable GPS System.</u> The High-Usability system consisted of a touch screen and suction cup mounting bracket that was attached to a painted (black) aluminum plate, which was mounted by Velcro to the dash. The plate covered the factory installed navigation system, as shown in Figure 19. A plastic disc was attached to the center of the plate to facilitate the use of the suction cups to attach the device. This method of attachment provided consistency of device location across systems.

Figure 19. High-Usability Navigation System

When using the High-Usability navigation system, selections were made by pressing buttons on the touch screen. This system had no additional physical buttons. If an error was made, the participant used the 'BACK' button, located in the lower left corner of each screen, to return to a previous screen. The participant could press the 'BACK' button repeatedly to return to the main menu, if desired.

The system occasionally had trouble with satellite reception in the laboratory. Thus, the participant sometimes experienced pop-up screens related to this problem. Participants were trained in how to respond to each of these error screens.

Low-Usability Portable GPS System. The Low-Usability system consisted of a touch screen and suction cup mounting bracket that was attached to a plate placed over the OEM navigation system, as shown in Figure 20.

Figure 20. Low-Usability Navigation System

When using the Low-Usability navigation system, selections were made by pressing the buttons on the touch screen. This system had no additional physical buttons. If an error was made and the participant wished to return to a previous screen, there was a ‘BACK’ button located in the lower left corner of each screen. This system also had a ‘SOURCE’ button located in the lower right corner of each screen to return to the main menu, if desired.

4.2.4 Procedure

Each participant completed one session, lasting approximately four hours. Upon arrival, the participant was escorted to a conference room and asked to read the Participant Information Summary, which described the experiment and set forth the terms of participation. The participants were encouraged to ask questions. After all questions were answered, the participant signed the documents, thereby giving informed consent to participate in the study. No individuals declined to participate.

The participant was escorted to the experimental vehicle and given an overview orientation of the vehicle controls and displays, including adjusting the seat and steering wheel. This was followed by an explanation of the monetary performance incentive system (Section4.2.7) and the Rating Scale Mental Effort (RSME, see Appendix D: Rating Scale Mental Effort (RSME)).

The participant was then asked to affix latex markers to his or her face for eye tracker calibration. This allowed the system to use facial features to help determine point of gaze and head position. During this procedure, the experimenter instructed the participant concerning head position and point of gaze. Eye tracker calibration was completed.

The participant was then given training and practice on the MDT. This was followed by driving task (car-following) training. Following training, the participant completed three practice drives. The first practice drive familiarized the participant with the simulator, its controls and the MDT. The second practice drive was car-following familiarization, in which there was no MDT present. For the third practice drive, the participant performed car following and the MDT concurrently.

Data collection began following a break. The experimenters, who were seated at a control center behind the vehicle, were able to communicate directly with the participant using a speaker and microphone system.

Experiment 3 required participants to complete eight main test trials. The first and last main test trials were baseline trials, in which the participants performed car following plus MDT, but no navigation system task. After the first baseline trial, navigation system task training began with an overview that pertained to all three navigation systems used in the experiment. Participants then worked with each of the three navigation systems in an order determined by the experimental design. For each system, participants received training, stationary practice, practice while driving (car following plus MDT) and then the main trial for each of the two task types (ADD and PD) using the first navigation system. Participants were offered a break following the completion of both tasks on each system. This process was repeated for the other two navigation systems.

Immediately after every trial (practice and main trials), the participant completed the RSME and was given performance feedback. If a participant took a break when offered, the experimenters would stop the eye tracker logging and restart it upon the participant's return to the vehicle.

At the completion of data collection, the participant was escorted to the conference room. In the conference room, the participant completed the simulator sickness questionnaire, received a copy of the Participant Information Summary form and received payment for participation plus performance incentives. The experimenters answered any questions and thanked the participant for his or her participation.

4.2.5 Driving Tasks

STISIM. The car-following task was the same task that was described in Section 2.3.1 Based on the results of Experiment 1, we used the 'moderate' speed signal, shown in Figure 1.

Multiple Target Detection Task (MDT). Based also on the results of Experiment 1, car-following was always performed together with the MDT, in which drivers responded to a sequence of targets, presented one at a time at one of the locations shown in Figure 3.

4.2.6 Secondary Tasks

Secondary tasks were identical to those involving navigation systems that were used in Experiments 1 and 2. They included destination entry by address (ADD) and selecting a previous destination (PD). The ADD task was self-paced; participants were required to enter a sequence of addresses, each of which consisted of a state, city, street name, and house number.

Entries were made using a virtual touch-screen keyboard, which required the participant to spell out the beginning letters, until a list was automatically generated by the system. At this point, the list was searched and the correct entry selected. The PD task was also self-paced; participants were required to find a sequence of addresses that had previously been entered into the respective systems. Participants activated the system list of previous destinations and then scrolled through this list to find the specified destination. Performance of these tasks differed slightly for each system. A comparison among the three systems is presented Appendix F. Pictures of the display screens for each task type by navigation system can be found in Appendix G: Display Screen Images of 3 Navigation Systems, Experiment 3

4.2.7 Monetary Incentives

In addition to a base pay of $26 per hour, participants had the opportunity to earn a modest amount of additional money during the experiment. The actual amount of money awarded per trial was based on participants' performance in the three tasks shown in Table 18. Incentive amounts were established to reflect the following priorities: (1) Car-following was the most important task; (2) in-vehicle navigation system tasks were of secondary importance; and (3) the target-detection task had lowest priority.

Table 18. Experiment 3 Incentive Amounts per Trial

	Performance		
Task	**Good**	**Acceptable**	**Poor**
Car Following Task	$1.80	$0.90	$0.0
Navigation System Task	$1.40	$0.70	$0.0
Target-Detection Task	$0.80	$0.40	$0.0
Total	$4.00	$2.00	$0.0

4.2.8 Data Reduction

Data were reduced to obtain the same measures that were used in Experiment 1. These included: car-following coherence (Cohere (CC)), car-following delay (Delay (CC)), mean headway (M Hdwy), standard deviation of headway (Std Hdwy), standard deviation of lane position (SDLP), steering entropy (Steer Entropy), MDT mean response time (MDT MRT), MDT proportion of correct responses (MDT P Corr), percentage of time viewing road center (PRC), the proportion of long glances (P Long Glance), and the RSME subjective workload ratings.

4.3 Results

We used Proc Mixed of SAS (Version 9.1.3) to compute an analysis of variance (ANOVA) for each dependent measure. The first set of analyses examined differences between the secondary task conditions collapsed across systems and baseline trials. This set of analyses was intended to answer two specific questions: (1) whether metrics showed differences between baseline trials (No secondary task) and trials involving navigation tasks, and (2) whether address entry tasks (ADD) were more disruptive than previous destination selection tasks (PD), generally.

For these analyses, all metrics revealed strong and statistically significant differences between baseline trials and trials with secondary task conditions. Four metrics were found to differentiate between the ADD and PD tasks. Results of statistical tests for metrics that exhibited significant or marginally significant differences between the task conditions are summarized in Table 19. Means for the task conditions are presented in Figure 21.

Table 19. Metrics Exhibiting Consistent Differences between Navigation Task Conditions (Experiment 3)

Metric	**t test result**	**Statistical test significance**
Cohere (CC)	t(250) = -1.70	p = .09+
RSME	t(250) = 5.03	p < .0001*
MDT MRT	t(250) = 2.27	p = .02*
P Long Glance	t(143) = 1.85	p = .07+

* Statistically significant difference ($p < .05$)
+ Marginally significant ($.05 < p < .10$)

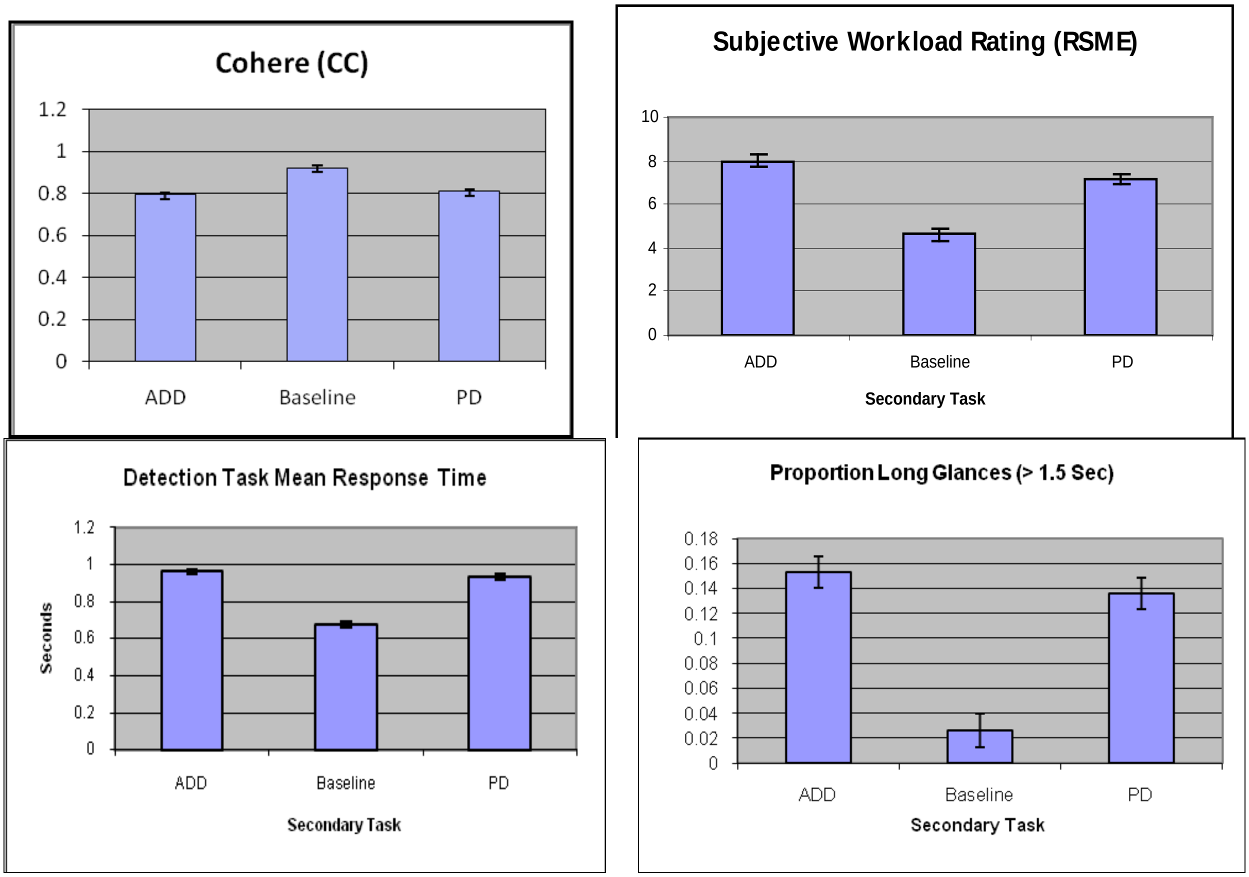


Figure 21. Task Condition Means (± Standard error) – Experiment 3

Differences shown in Figure 21 are all in the expected direction, indicating, for example that relative to PD tasks, the ADD tasks were associated with slightly lower car-following coherence, slightly longer detection task response times, a slightly higher proportion of long glances, and more subjective workload.

The second set of analyses eliminated baseline trials and examined effects of System (OEM, High-Usability [HiU], Low-Usability [LoU]), Tasks (ADD, PD) and their interactions. These results are summarized in Table 20. Table entries in the three middle columns are probability values associated with statistical tests. Thus, p values greater than 0.10 indicate no statistically significant differences; p values less than 0.05 indicate statistically significant differences; differences with p values between 0.05 and 0.10 are considered marginally significant and interpreted as potentially meaningful. Interpretations are presented for each statistically or marginally significant effect and significant interaction effects are presented in Figure 25.

Table 20. Summary of Main and Interaction Effects for Each Performance Measure (Experiment 3)

Measure	System (1)	Task (2)	Interaction (3)	Interpretation
Delay (CC)	0.44	.056+	.38	(2) ADD > PD
Coherence (CC)	.70	.17	.38	
Steer Entropy	.21	.47	.11	
MDT Mean RT (MDT MRT)	.02*	.03*	.62	(1) (HiU= LoU) > OEM (2) ADD > PD,
MDT P Correct	.37	.16	.05*	(3) Figure 22
SDLP	.26	.36	.01*	(3) Figure 22
P Road Center (PRC)	.42	.09+	.53	(2) PD > ADD
P Long Glance	<.0001*	.04*	.004*	(1) OEM< HiU < LoU (2) ADD > PD (3) Figure 22
RSME	.0015*	<.0001*	.39	(1) (OEM = HiU) < LoU (2) ADD > PD

* Statistically significant difference ($p < .05$)
\+ Marginally significant ($.05 < p < .10$)

RSME ratings, which represent subjective workload assessments, have been used throughout this work to provide expectations concerning the patterns of results for the objective measures. In Experiment 3, as shown in Table 20, RSME was associated with a significant Task main effect (Task column: $p < .0001$); specifically participants rated the ADD task to be more demanding than the PD task (Interpretation column: ADD > PD). Participants also rated the individual systems differently (System column: $p = .0015$); specifically, the Low-Usability system was rated as more demanding than the other two systems, which were not rated significantly different from one another (Interpretation column: ([OEM = HiU] < LoU).

Among the objective measures, the proportion of long glances (P Long Glance) and the Multiple Detection Task Mean Response Time (MDT MRT) provided the most comparable pattern of results. The Task main effects for both of these measures are in the predicted direction; specifically, the ADD task was associated with longer detection task response times and higher proportions of longer glances than the previous destination (PD) task. Car following delay (Delay (CC)) was associated with a marginally significant Task main effect, which has a slightly higher probability of being due to chance; however, the results indicate that the ADD task had slightly longer delays than the PD task. The eye-position based measure of proportion road center (PRC) also revealed a marginally significant Task main effect, reflecting the tendency of drivers to spend a slightly greater proportion of their driving time looking away from the road center in the ADD task than in the PD task.

Significant interactions were identified for three of the measures (MDT P Correct, SDLP, P Long Glance), representing more complex differences between the combinations of Systems and Tasks. These effects are presented in Figure 22.

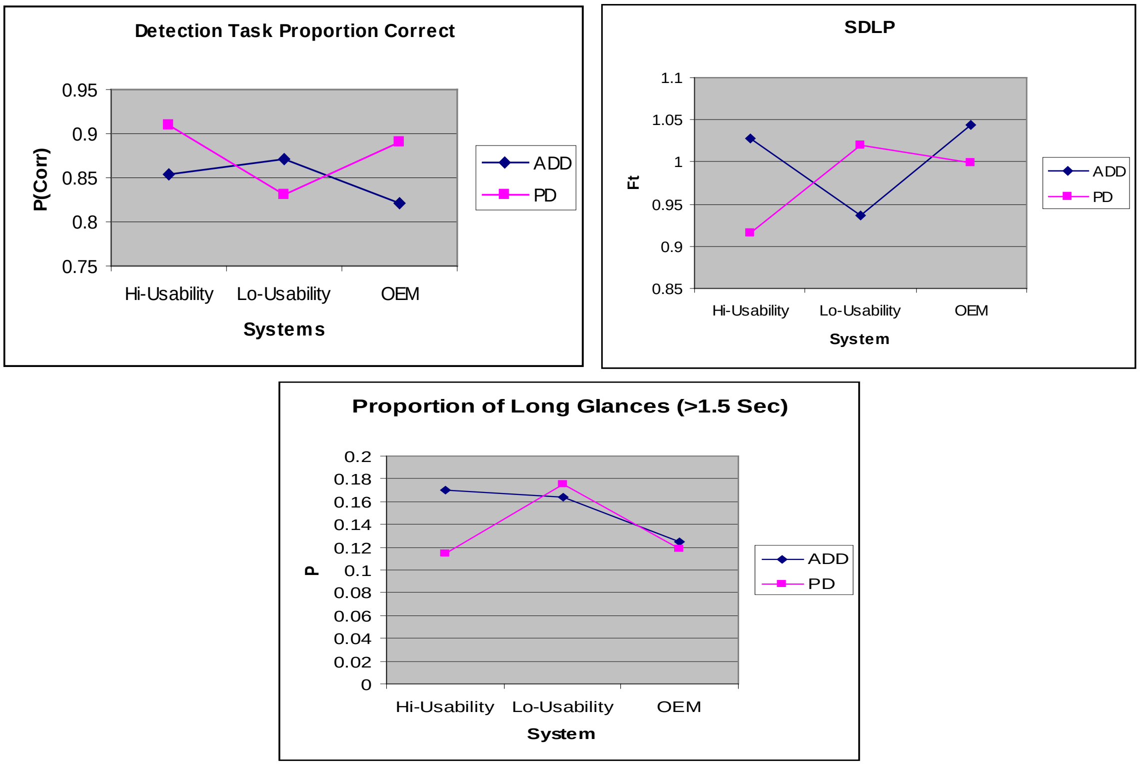


Figure 22. Significant Task by System Interaction Effects – Experiment 3

For the proportion of targets correctly detected (MDT P Correct), the results were consistent with predictions for the OEM and High-Usability system, but not for the Low-Usability system. Specifically, for the former two systems, the PD tasks were associated with higher proportions of targets detected than the ADD tasks, reflecting the higher level of demand associated with the ADD tasks. This pattern was reversed for the Low-Usability system, for which the ADD task had slightly elevated proportions of targets detected.

A similar pattern was evident for the SDLP measure, as shown in Figure 22. Although the System main effect was not statistically significant, the pattern suggests that the OEM system was associated with generally higher SDLP values. The OEM and High-Usability systems exhibited trends consistent with predictions; SDLP values were greater for the ADD task than for the PD task. This pattern was reversed for the Low-Usability system, for which the ADD task decreased SDLP relative to the PD task. The difference between the tasks for the High-Usability system was larger than for the OEM system.

4.4 Discussion

Differences between baseline trials and trials involving secondary tasks were strong for most metrics, reflecting the significant degradation of driving performance associated with performing secondary tasks using navigation systems. Although the differences between the two navigation system tasks (ADD and PD) were generally consistent with the differences between these tasks observed in Experiment 1, the effects were weaker with fewer differences reaching statistical significance. These differences reflect the fact that in Experiment 3 this test was made using combined data from three systems while the test in Experiment 1 was made using data from one system. The examination of the significant Task x System interactions revealed the nature of the differences between the three systems. For SDLP, the absence of a significant Task main effect was due to the fact that the Low-Usability system exhibited a pattern that was contrary to the other two systems. Specifically, for the Low-Usability system, the PD task was associated with a higher level of SDLP, indicative of higher performance degradation, than the ADD task. This pattern was the opposite of that observed for the other two systems and in our previous work in which the ADD task consistently was associated with higher levels of performance degradation than the PD task. A similarly inconsistent pattern of results was observed for the Proportion of Detection Targets correctly detected. For this metric, the Low-Usability system was associated with relatively fewer targets detected in the PD condition than in the ADD condition. We had expected fewer targets to be detected in the ADD condition, which is more demanding than the PD task. This pattern was observed for the other two systems. While the presence of significant Task x System interactions for some metrics helps explain the weaker observed differences in Experiment 3, the reasons for these interactions are not readily apparent based on the analyses conducted.

5.0 DISCUSSION

5.1 Experiment 1

Experiment 1 used three categories of secondary tasks, including an auditory/vocal (N-back) task for which the load was primarily cognitive, and two visual-manual tasks. The navigation system tasks were complex visual-manual tasks, requiring a number of different steps, alternating between using keyboards, searching through lists and identifying targets. The Circles task was a visual search task, which required a relatively simple visual discrimination (which was more difficult in this experiment than in our previous work) and a manual response. In Experiment 1, all of the metrics exhibited strong differences between the two categories of visual-manual tasks. Similarly, with the exception of the subjective workload ratings (RSME), all metrics exhibited strong differences between the N-back task and the navigation system tasks, with the latter category being associated with consistently higher levels of performance degradation, relative to the former. Interestingly, this was the first evidence that subjective workload ratings did not correspond to the differences observed among objective measures. Thus, while all objective measures revealed greater performance degradation for the navigation system tasks relative to the N-back tasks, the participants' ratings of workload demands were not statistically different. This difference trended in the predicted direction; however, the difference was too small to be statistically significant. This was our first use of the N-back task and feedback from participants indicated that it was mentally demanding, particularly in the 2-back condition. The fact that the instructions for the subjective ratings focused on mental effort while the objective metrics were generally more sensitive to visual and manual distraction effects than to mental or cognitive demands may have contributed to this weaker trend.

Differences between the simple Circles task and the N-back task indicated that some metrics are sensitive to both visual-manual and cognitive interference. Specifically, both categories of tasks interfered with car-following and detection task performance. A possible explanation for the car-following effects is that car-following has both a visual and cognitive component. Similarly, detection task performance is sensitive both to visual loads that require drivers to divert their eyes from the forward scene and cognitive load which causes drivers to fail to notice targets in the visual field. In contrast, measures of lateral vehicle control (SDLP, Steer entropy) and visual behavior (PRC, P Long Glance) revealed different patterns of degradation between these two task categories. Specifically, the simple visual manual task was associated with more lane position variability, more steering error, less time looking ahead, and proportionately more long glances than the auditory/vocal task.

With respect to the differences between conditions within secondary task categories, metrics were generally able to differentiate between conditions in the N-back task and in the navigation system tasks, but not for the Circles task. Specifically, both car following measures (Delay (CC), Cohere (CC), and Std Hdwy) and detection task measures (MDT MRT and MDT PCorr) revealed differences between 1-Back and 2-Back conditions, indicating sensitivity for distinguishing between different cognitive loads. The same pattern was generally found for the navigation system tasks, with several exceptions; first, car-following differences were not as consistent across all measures (Std Hdwy was not sensitive to these differences), and only one of the two detection task measures was sensitive. SDLP did exhibit differences between the two navigation system tasks, but not between the cognitive tasks. Neither of the eye-position based measures was sensitive to differences between any of the three task categories. Finally, none of the metrics differentiated between the two levels of demand used in the Circles Task. As indicated previously, this was due to our decision to use a relatively small difference between conditions, unlike our previous work, which had used a much larger difference and found consistent sensitivity among most metrics. The absence of RSME differences between Circles task conditions indicates comparable sensitivity between subjective workload assessments and objective metrics.

5.2 Experiment 2

The second experiment adopted a different approach to the assessment of distraction effects. In particular, the driving-simulator-based test used in Experiment 1 has been developed focusing exclusively on the amount of performance degradation associated with different secondary task conditions. Participants performed secondary tasks repeatedly over a predefined time interval; summary performance measures were computed to estimate the average level of degradation occurring at any point during this interval. In this approach, no attempt is made to incorporate the duration of the task into the assessment. Thus, while the estimate of the average level of performance degradation represents the distraction potential, a more complete estimate of the overall exposure to risk associated with a particular secondary task requires combining this estimate with the expected duration of the task. The occlusion paradigm provides an estimate of the task duration, obtained under conditions of interrupted performance as in driving. Unfortunately, because participants typically have no primary task load (to simulate the demands of driving), the estimates of the time required to complete a task under conditions of occlusion are contaminated by the fact that participants can continue to work on the secondary task during occluded intervals (blind operation), which are not counted as part of the task duration. The Enhanced Occlusion Task (EOT) addresses this concern by adding an auditory tracking task, intended to simulate the demands of driving without interfering directly with the visual allocation of attention required by the occlusion protocol. The objectives of Experiment 2 were to determine the extent to which blind operation was eliminated by the EOT and to determine whether the EOT improved the sensitivity of the derived R metric, which represents task resumability, relative to the traditional occlusion protocol. The same two navigation system tasks, including destination entry by address (ADD) and selecting a previous destination (PD) were used in Experiment 2.

Generally, we found that the EOT eliminated part of the blind operation, but not all of it. Specifically, the traditional occlusion paradigm provided task duration estimates what were on average 77 percent of the task performance time obtained under a static condition (continuous performance). If it assumed that there is no time cost for switching in the occlusion protocol, this implies that 23 percent of the effort required to perform the task was accomplished during occluded intervals. With the EOT, the corresponding percentage was 89 percent, with 11 percent of the effort accomplished during the occluded intervals. We also found that the R metrics computed with data from the EOT were significantly different from those computed with the data obtained from the traditional occlusion paradigm, but were no more sensitive to hypothesized differences in task resumability between the navigation task conditions. Unfortunately, this was a rather weak test as there has been no independent confirmation that the ADD tasks were less resumable than the PD tasks, despite the well-documented differences in workload and in the associated amounts of performance degradation. RSME measures indicated that the two tasks differed in terms of subjective workload estimates.

The EOT provided an additional data source, namely measures of auditory tracking performance. It was found that the ADD task was associated with a significantly higher level of tracking error (AMD and P RAP) than the PD task. This finding is consistent with the simulator test results and implies that estimates of task duration obtained with the EOT paradigm must be considered together with the primary task measures. It appears that task resumability is not related to the level of primary (driving) task performance degradation associated with secondary task performance. Finally, because the task duration estimates provided by the EOT paradigm differ from those provided by traditional occlusion, a stronger connection between estimates based on occlusion and estimates for identical tasks obtained in driving situations is needed to establish the relation between TSOT and real-world task duration.

5.3 Experiment 3

Experiment 3 utilized the distraction potential simulator-based test protocol following modifications based on the results of Experiment 1. Modifications included use of the multiple-target detection task (MDT) and the moderate (lesser) level of car-following task difficulty. The effects of three different navigation systems were compared. Participants performed the two navigation system tasks (ADD and PD) using one OEM system and two portable systems, which differed in their rated usability. It was hypothesized that the OEM product, by virtue of its design to be used specifically in the driving context, would be less potentially distracting than either of the portable systems and that the High-Usability system would be less potentially distracting than the Low-Usability system. Metrics revealed strong and consistent differences between baseline driving and driving with a secondary task. Three objective metrics (car-following coherence, detection task mean response time and the proportion of long glances) revealed differences between the ADD and PD tasks generally; however these differences were weaker than those observed in Experiment 1, reflecting the fact that the effects of the three systems on driving performance were not consistent. Additional analyses were conducted to explore these differences. Most notable among the driving performance metrics was the apparent lack of sensitivity between task conditions for the SDLP metric, which had been among the strongest metrics for differentiating between task conditions for visual-manual tasks. This was due to a significant interaction between Systems and Tasks, which revealed that the predicted differences between tasks were evident for the OEM and Hi-Usability systems but reversed for the Lo-Usability system. A similar pattern was observed for the detection task proportion of correct responses. The explanation for this pattern of results is not readily apparent; however, the presence of significant System x Task interactions suggests that multiple tasks should be performed with each system as part of any assessment of the distraction potential of IVIS technologies. Similarly, conclusions about distraction potential for specific tasks cannot be made without consideration of the specific device used.

5.4 Practical Significance of Results

While many differences between task conditions in this research were found to be statistically significant, it is fair to ask whether these differences are meaningful or practically significant. For example, does a statistically significant increase in lane position variability or in target detection response time have a meaningful impact on safety? Unfortunately, due to the wide variation in real-world driving conditions these are very difficult questions to answer. Driving alone on a multi-lane highway allows more room for lateral positioning error without an increased crash risk than driving in a congested tunnel with narrow lanes or concrete barriers immediately adjacent to the travel lane. For this reason, it is virtually impossible to make general statements about the meaning of a difference of a given magnitude and whether such a difference represents a meaningful increase in crash risk.

A related question of practical importance in the context of driver distraction involves determining the acceptable level of driving performance degradation associated with the concurrent performance of a given secondary task. The use of benchmarks has evolved as one way to address this question. Early work in the development of distraction metrics, performed under the auspices of the HASTE and ADAM, used both abstract laboratory tasks and real in-vehicle tasks for calibration purposes. Metrics that could reliably detect pre-established differences between conditions were considered sufficiently sensitive to the effects of distraction. Unfortunately, there was no strong safety-related basis for the differences considered to be meaningful. Rather, the selection of meaningful differences was based on consensus among participating organizations. More recently, AAM Guidelines identified radio tuning as a benchmark task. Accordingly, secondary tasks associated with significantly greater levels of performance degradation than radio tuning are considered unacceptable for use while the vehicle is moving. Auto manufacturers have also begun to lock out destination entry by address in a moving vehicle, which indicates an emerging consensus that this task is too demanding to be permitted while the vehicle is in motion. It follows that a task must be associated with significantly less performance degradation to be considered acceptable for use when driving.

The use of benchmarks represents an improvement over the use of differences alone; however, it still relies on statistical significance, which can be manipulated by increasing the sample size. Specifically, a difference of a given magnitude is more likely to be determined to be statistically significant with a larger sample size. This problem has been addressed by the emerging consensus concerning the number of participants required for an experimental evaluation. Specifically, the state-of-the-practice is to conduct relatively small-scale experiments (using approximately 20 participants) and statistically compare the performance degradation for secondary tasks on selected metrics with the degradation on the chosen benchmark tasks. Note that this approach is only suitable for a test protocol that bases decisions on the existence rather than the absence of statistically significant differences.

In the absence of strong ties to safety and the considerable difficulty associated with determining whether a particular secondary task is more dangerous to perform while driving than another task, researchers and auto manufacturers have relied on benchmarks and consensus concerning the details of the experimental protocols and the specific performance metrics used to address such questions. The present work was conducted within this framework and is thus subject to the same limitations.

6.0 CONCLUSIONS

1. The development and implementation of a driving-simulator-based test that can be used to assess the distraction potential of secondary tasks performed with OEM equipment in production vehicles or portable devices is feasible. The test can be implemented without requiring significant setup and without damaging vehicles.
2. Core metrics include those sensitive to visual-manual distraction (standard deviation of lane position, car-following delay, and detection task response time) and those sensitive to cognitive distraction (car-following delay, detection task response time, and detection task proportion of correct responses).
3. The Enhanced Occlusion Technique (EOT) eliminates part but not all of the blind operation inherent in the occlusion technique and thus improves the validity of task duration estimates; however task duration estimates obtained with the EOT require a stronger connection to comparable values obtained in a controlled driving situation.
4. The simulator-based test of distraction potential focuses on the dynamics of distraction rather than the duration of the distracting activity. Estimates of distraction potential can be combined with task duration estimates provided by the EOT to compute estimates of drivers' exposure to risk.
5. Performance of the two navigation tasks that have previously shown different effects on driving performance did not have consistent effects on driving performance when performed with different systems. Conclusions about the distraction potential of a particular task cannot therefore be made without consideration of the system on which the task was performed.
6. Distraction potential test results were not consistent with usability ratings, reflecting the possibility that usability ratings may not correlate strongly with distraction potential effects observed in a driving situation, in which tasks are performed intermittently.
7. Benchmark tasks are needed to establish levels of acceptable doses of distraction. The N-back task was consistently disruptive to driving performance. The 2-back condition could thus serve as a starting point for setting a limit for acceptable "dose" of cognitive distraction.

7.0 REFERENCES

Angell, L. S., Auflick, J. L., Austria, P. A., Kochhar, D. S., Tijerina, L., Biever, W. J., Diptiman, T., Hogsett, J. R., Jr., & Kiger, S. M. (2005). Driver Workload Metrics Project, Task 2 CAMP.

Barickman, F. S. & Goodman, M. J. (1999). Micro DAS: In-vehicle portable data acquisition system. Transportation Research Record, 1689, 1-8.

Boer, E. R. (2000). Behavioral entropy as an index of workload. In Proceedings of the IEA 2000/HFES 2000 Congress (pp. 3-125-3-128).

Briem, V. & Hedman, L. R. (1995). Behavioural effects of mobile telephone use during simulated driving. Ergonomics, 38, 2536-2562.

Brookhuis, K. A., De Vries, G., & De Waard, D. (1991). The effects of mobile telephoning on driving performance. Accident Analysis and Prevention, 23, 309-316.

Brookhuis, K. A., Waard, D. d., & Mulder, B. (1994). Measuring driving performance by car-following in traffic. Ergonomics, 37, 427-434.

Carsten, O. & Brookhuis, K. (2005a). Issues arising from the HASTE experiments. Transportation Research Part F: Traffic Psychology and Behaviour, 8F, 191-196.

Carsten, O., Merat, N., Janssen, W. H., Johansson, E., Fowkes, M., & Brookhuis, K. (2005b). Human Machine Interaction and the Safety of Traffic in Europe (Rep. No. HASTE Final Report 1.0). University of Leeds: Institute for Transport Studies.

Consumer Reports (2009). http://www.consumerreports.org/cro/cars/tires-auto-parts/gps/gps-ratings/ratings-overview.htm

Gelau, C. & Schindhelm, R. (2009). Enhancing the occlusion technique as as assessment tool for driver visual distraction. Presented at the ITS World Congress, Stockholm, Sweden (Sept. 21 – 25).

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: Traffic Psychology and Behaviour, 6, 23-36.

Horberry, T., Stevens, A. Cotter, S, Robbins, R. & Burnett, G. (2007). Development of an occlusion protocol with design limits for assessing driver visual demand. Transport Research Laboratory (Report PPR 256). Wokingham, Berks, UK: Transport Research Laboratory.

ISO/TC 22, S. 1. Road vehicles – Ergonomic aspects of transport information and control systems — Simulated lane change test to assess in-vehicle secondary task demand. ISO/WD 26022, 1-31. 2004. International Organization for Standardization.
Ref Type: Generic

ISO (2007). Road vehicles – Ergonomic aspects of transport information and control systems – Occlusion method to assess visual demand due to the use of in-vehicle systems International Standard 16673. Geneva, Switzerland: International Organization for Standardization.

Jamson, A. H. & Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. Transportation Research Part F, F8, 79-96.

Klatzky, R.L., Giudice, N.A., Marston, J.R., Tietz, J., Golledge, R.G., & Loomis, J.M. (2008). An n-back task using vibrotactile stimulation with comparison to an auditory analogue. Behavior Research Methods, 40 (1), 367-372.

Mattes, S. (2003). The lane change task as a tool for driver distraction evaluation. In GfA/17th Annual Conference of the International-Society-for-Occupational-Ergonomics-and-Safety (ISOES) Stuttgart, Germany: Ergonomia Verlag OHG, Bruno-Jacoby-Weg 11, D-70597.

Merat, N. & Jamson, A. H. (2007). Multisensory signal detection: A tool for assessing driver workload during IVIS management. Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, Washington.

Merat, N., Johansson, E., Engstrom, J. A., Chin, E., Nathan, F., & Victor, T. W. (2007). Specification of a secondary task to be used in safety assessment of IVIS (Rep. No. IST-1-507674-IP). Adaptive Integrated Driver-Vehicle Interface.

Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2004). Task interruptability and duration as measures of visual distraction. Applied Ergonomics, 35, 207-213.

Pettitt , M. A., Burnett, G. E., Bayer, S., & Stevens, A. (2006). Assessment of the occlusion technique as a means for evaluating the distraction potential of driver support systems. IEEE Proceedings Intelligent Transportation Systems, 153, 259-266.

Ranney, T. A., Baldwin, G. H. S., Vasko, S. M., & Mazzae, E. N. (2009). Attentional Demands of Operating In-Vehicle Devices: Establishing Best Practices (Rep. No. DOT HS 811 231). Washington, D.C.: U.S. Department of Transportation: National Highway Traffic Safety Administration.

Ranney, T. A., Harbluk, J. L., & Noy, Y. I. (2005). Effects of voice technology on test track driving performance: Implications for driver distraction. Human Factors, 47, 439-454.

Ranney, T. A., Mazzae, E., Baldwin, G. H. S., & Salaani, K. (2007). Characteristics of voice-based interfaces for in-vehicle systems and their effects on driving performance (Rep. No. DOT HS 810 867). Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Recarte, M. A. & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination and decision making. Journal of Experimental Psychology: Applied, 9, 119-137.

Reimer, B. (2009). Cognitive task complexity and the impact on drivers' visual tunneling. (Paper 09-2953). DVD Compendium of Papers Transportation Research Board 88th Annual Meeting, January 11-15.

Schindhelm, R. & Gelau, C. (2008). Enhanced Occlusion Technique (EOT): Assessment if IVIS with regard to driver distraction. Adaptive Integrated Driver-vehicle Interface (AIDE). Final workshop and exhibition, Gothenburg, Sweden, April 15-16..

Stevens, A., Bygrave, S., Brook-Carter, N., & Luke, T. (2004). Occlusion as a technique for measuring In-Vehicle Information System (IVIS) visual distraction: a research literature review (Rep. No. TRL609). Berkshire, United Kingdom: TRL Limited.

Victor, T. W., Engstrom, J. A., & Harbluk, J. L. (2008). Distraction Assessment Methods Based on Visual Behaviour and Event Detection. In M.A.Regan, J. D. Lee, & K. L. Young (Eds.), Driver Distraction: Theory, Effects and Mitigation, London: CRC Press).

Victor, T. W., Harbluk, J. L., & Engstrom, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. Transportation Research Part F, 8F, 167-190.

Westfall, P. H., Tobias, R. D., Rom, D., Wolfinger, R. D., & Hochberg, Y. (2003). Multiple Comparisons and Multiple Tests Using SAS. Cary, North Carolina: SAS Institute Inc.

8.0 APPENDICES

8.1 Appendix A: Analysis of FaceLAB Data

The FaceLAB data consists of a set of two measurements describing gaze location. These two measurements are called "yaw" and "pitch". The "yaw" values describe a right to left angular displacement of the gaze while "pitch" describes an up and down angular displacement of the gaze.

Often, FaceLAB measurements are difficult to interpret because there can be unexplained offsets in the data. These offsets are time varying, but they seem to be of a fairly low frequency. Thus, a cluster of gaze measurements tends to describe a single region where the subject was looking.

Figure 23 shows a cloud of yaw, pitch pairs describing the subject's gaze location over a two and a half minute period, which represents one driving trial. The image in Figure 23 is the analysis report derived from Facelab data for a single following event, specifically subject 11 run 7. The intersection of the red coordinate axes occurs at a yaw of zero and a pitch of zero.

8.1.1 Finding the Primary Maximum Density Regions

The first step in analyzing the FaceLAB data is to scan the two dimensional yaw/pitch data with a circular window of a radius of 8 degrees. The scan step was 0.57 degrees. At each window location on the yaw/pitch plane, the number of points within the window is found and the window location with the largest number of points is recorded as the primary maximum density region. Figure 23 shows the primary maximum density region for subject 11 run 7 in blue (left-most circle).

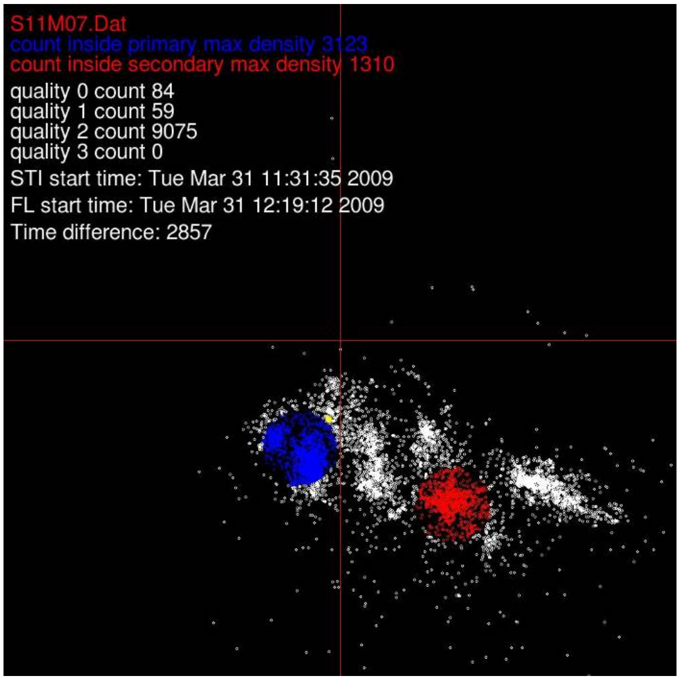


Figure 23. Cloud of Yaw, Pitch Pairs Describing Gaze Location

The text in the analysis report image shown in Figure 23 indicates that there are 3123 points within the maximum density region. The other text indicates how many points of each of the four quality levels there are in the data set. Only points of quality 2 or above are used to find the maximum density regions. In all these experiments, the wider view Facelab operating mode was used. This wide view operating mode has quality 2 as the maximum quality value.

The other text printed on the image summary is generally for debugging purposes only. However, in the interest of completeness, the "STI start time" is the run start time for the simulator computer. The "FL start time" is the run start time for the FaceLAB computer. The difference between the two times is the time difference between each computer's clocks. Generally, the computer's clocks are not used to synchronize the data. However, occasionally, the computers' real time clocks were used to save data when an unexplained problem caused the Facelab frame number to fail to show up in the uDAS log.

8.1.2 Finding Secondary Maximum Density Regions

Many tasks required dividing the subject's attention. For this reason, it is necessary to discover a second maximum density region which represents another location where the subject's gaze lingers during the test run.

The procedure for finding the second maximum density region is similar to finding the primary region. The only difference is that points within the primary density region and within some buffer region surrounding the primary density region are removed from the window's scan, and these points are not considered in the clustering.

In prior versions of the algorithm for finding the secondary maximum density region, the possible secondary maximum density regions were allowed to be immediately adjacent to the primary maximum density region. However, if the secondary cluster appears too close to the primary region, it may not represent the subject looking away from the roadway at the secondary task. Instead, the secondary cluster could just be a collection of points which are still rather close to the primary maximum density region. Thus, the processing algorithm now incorporates a buffer which ensures that the secondary maximum density region is two region radii away from the primary cluster.

Figure 23 shows the secondary maximum density region in red (right-most circle). In this particular case, a significant, distinct secondary maximum density region appears in the data. In cases where the subject does not take his or her eyes off the road for significant amounts of time, the secondary maximum density region may not encompass many points and it may be in close proximity to the primary maximum density region. Figure 24 shows such a situation.

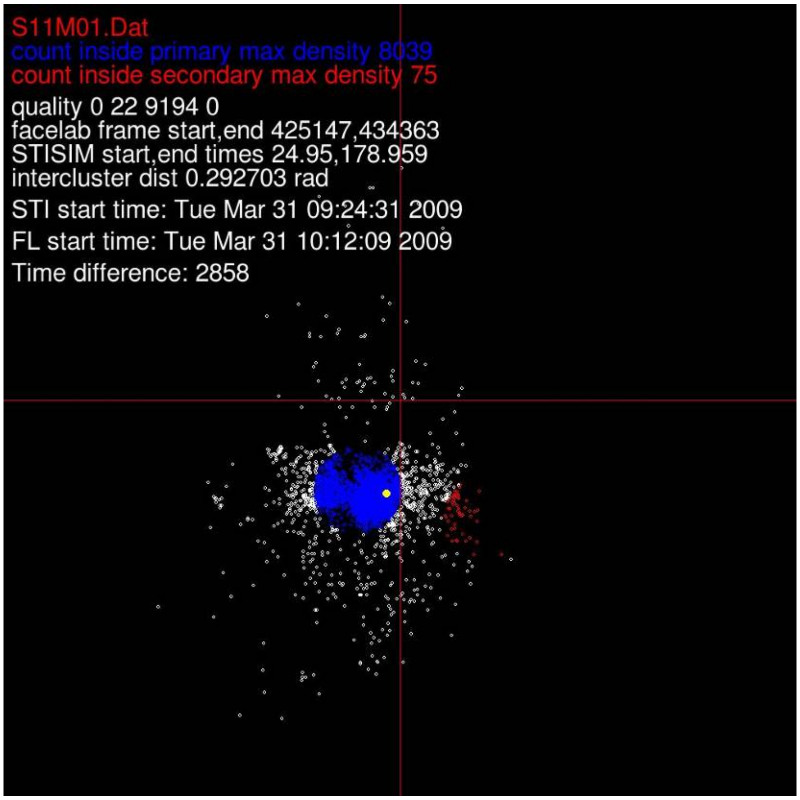


Figure 24. Example of Minimal Subject Eyes Off Road Time

The secondary maximum density region in Figure 24 is as close as the buffer around the primary maximum density region allows, which is no less than 2 density region radii. However, the number of samples located within the secondary region (75, i.e., less than 1 percent of 9194 total samples) indicates that the subject did not spend a significant amount of time looking at anything other than the road. This pattern is typically seen on trials without a designated visual-manual secondary task.

In some cases, the region with the most gaze points is not the primary maximum density region, but rather the secondary maximum density region. Essentially, this situation describes a subject whose gaze spends less time on the road than on a secondary task. In this case, whichever maximum density region is closest to the boxology centroid is designated as the primary maximum density region. “Boxologies” are described below in Section 8.1.6. Generally, boxologies are the outcome of a procedure intended to find empirical evidence of the location of the roadway center within the FaceLab yaw/pitch coordinate framework.

8.1.3 Ranking the Regions

The simplest method of ranking high density regions is to count the number of points within each region and compare that number with the total number of points. This number should give an idea of what percentage of time during the test run that the test subject is looking at objects represented by the primary and secondary maximum density regions.

The degree to which the maximum density regions are spread out may be found by calculating the deviation of the region. The centroid in yaw (ψ_c) and pitch (θ_c) for each point within the maximum density region may be found with:

$$\psi_c = \frac{\sum_i \psi_i}{n}$$ for each ith point of the n points within the maximum density region

$$\theta_c = \frac{\sum_i \theta_i}{n}$$ for each ith point of the n points within the maximum density region

The square root of the average squared distance:

$$\sigma = \sqrt{\frac{\sum_i (\psi_i - \psi_c)^2 + (\theta_i - \theta_c)^2}{n}}$$

is the standard deviation of the distance from the centroid.

8.1.4 Cluster Region Localization

The FaceLAB data does seem to have unpredictable offsets. For this reason, the location of the cluster representing the gaze associated with the secondary task is expressed in terms of the primary cluster's location. The location of this primary maximum density region is assumed to be the center of the rear end of the lead vehicle in the car-following scenario. Because the following task is performed only on straight road segments, this position is expected to vary only with significant changes in headway, i.e. the distance between the lead and subject vehicles.

One of the more straightforward methods of describing the location of the secondary cluster with respect to the primary cluster is by finding a distance and bearing from the primary cluster. One could use yaw/pitch pairs, which describe the locations of the primary and secondary clusters, (ψ_p, θ_p) and (ψ_s, θ_s), to find a distance and bearing (d, ϕ) within the two dimensional yaw, pitch data set:

$$d = \sqrt{(\psi_p - \psi_s)^2 + (\theta_p - \theta_s)^2}$$

$$\phi = \arctan 2(\theta_s - \theta_p, \psi_s - \psi_p)$$

8.1.5 Analyzing Transitions

In addition to the summary measures described above, it is possible to note when and for how long the gaze moves out of the primary region. For instance, the following table summarizes the excursions from the primary region for subject 7 trial 2.

Table 21. Characteristics of Excursions Outside the Primary Visual Cluster (Roadway Ahead)

Start Excursion	Time Outside Primary Cluster	Time Inside Secondary Cluster
0.08	0.18	0.00
0.58	0.77	0.00
1.43	0.08	0.00
1.58	0.13	0.00
1.77	0.27	0.00
2.75	2.15	1.60
5.23	1.47	0.78
6.77	0.15	0.00
7.80	1.17	0.72
9.83	1.40	1.35
11.83	1.30	0.95
14.38	1.78	1.67
17.67	1.55	1.48
20.98	1.28	1.08
23.02	1.60	1.50
24.72	0.10	0.00
25.57	0.72	0.42
27.77	0.27	0.18
28.07	0.22	0.00
34.07	0.17	0.00
34.47	0.02	0.00
34.53	0.57	0.00
35.43	0.18	0.00
35.83	0.35	0.00
38.20	0.20	0.02
39.15	0.78	0.42
40.70	1.38	1.18
60.18	0.10	0.02
61.97	0.18	0.03
64.48	0.30	0.03
80.80	0.48	0.12
86.48	0.17	0.00
86.85	0.07	0.00
87.00	0.13	0.00
99.07	0.22	0.03
103.15	0.40	0.05
103.60	0.02	0.00
103.75	0.07	0.00
105.50	0.02	0.02
110.50	0.87	0.70
114.15	0.20	0.00
114.38	0.25	0.00

All numbers in the above table are in seconds. The first column indicates the time within the follow event when the gaze excursion began. The second column indicates the duration of the excursion. The third column indicates the amount of time during this excursion that the gaze was inside the secondary cluster. If the gaze excursion did not enter the secondary cluster, the "time inside secondary" is zero.

8.1.6 Boxologies

In an effort to inject some measure of ground truth into the analysis of FaceLab's output, the experimenter instructs the subject to follow a cursor on the computer's projection screen with his or her gaze. The cursor then automatically moves through the corners and center of a rectangle. The resulting plot of the yaw and pitch measurements may be found in Figure 25. This ground truth procedure came to be referred to as a "boxology".

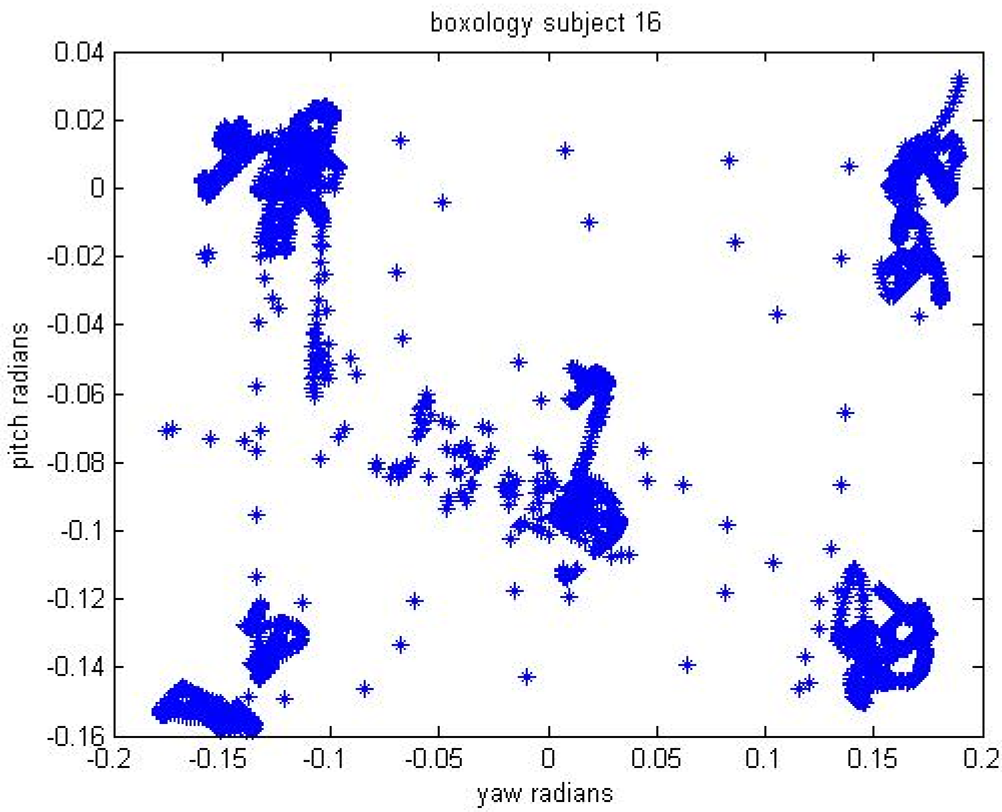


Figure 25. Plot of a Boxology

Figure 25 shows the distinctive "boxology" FaceLAB output where the five gaze clusters form a figure similar to the 5 side of a die. Four clusters occupy the corners of a rectangle and one cluster occupies the middle of the rectangle. The purpose of the boxology is to give some indication within FaceLAB's coordinate space of where the road lies. The centroid of this boxology is found and it is used to indicate which maximum density region is likely to be the road. The boxology centroid for the analysis report in Figure 23 is indicated by a small yellow dot.

8.1.7 Yaw and Pitch Measurements as Distance on the Screen

The yaw and pitch measurements may be transformed into distances measured on the projection screen. First, the yaw and pitch measurements are transformed into a three dimensional unit vector $\vec{a}$ pointing in the desired direction by using the following equations found in the FaceLAB manual (Seeing Machines, 2005):

$$\vec{a} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\sin(\psi)\cos(\theta) \\ \sin(\theta) \\ -\cos(\psi)\cos(\theta) \end{pmatrix}$$

If the screen is some distance d from the driver's head, $\vec{a}$ may be scaled such that it's component along the z-axis is equal to d. Thus, $\vec{a}$ becomes:

$$\vec{a}_d = \vec{a}\frac{d}{-\cos(\psi)\cos(\theta)}$$

where the x and y components of $\vec{a}_d$ are in whatever units were used to measure d, and the z component of $\vec{a}_d$ is constant and d.

8.1.8 Boxology Example in Terms of Distance

Figure 26 shows subject 12 boxology data transformed into inches by the procedure described above. The x-axis describes right to left coordinates in inches. The y-axis describes up and down coordinates in inches.

The distance from the subject to the screen is 190 inches. The on-screen length of the rectangle described by the boxology cursor was measured to be 51.25 inches while the on-screen height of the boxology rectangle was found to be 27.5 inches.

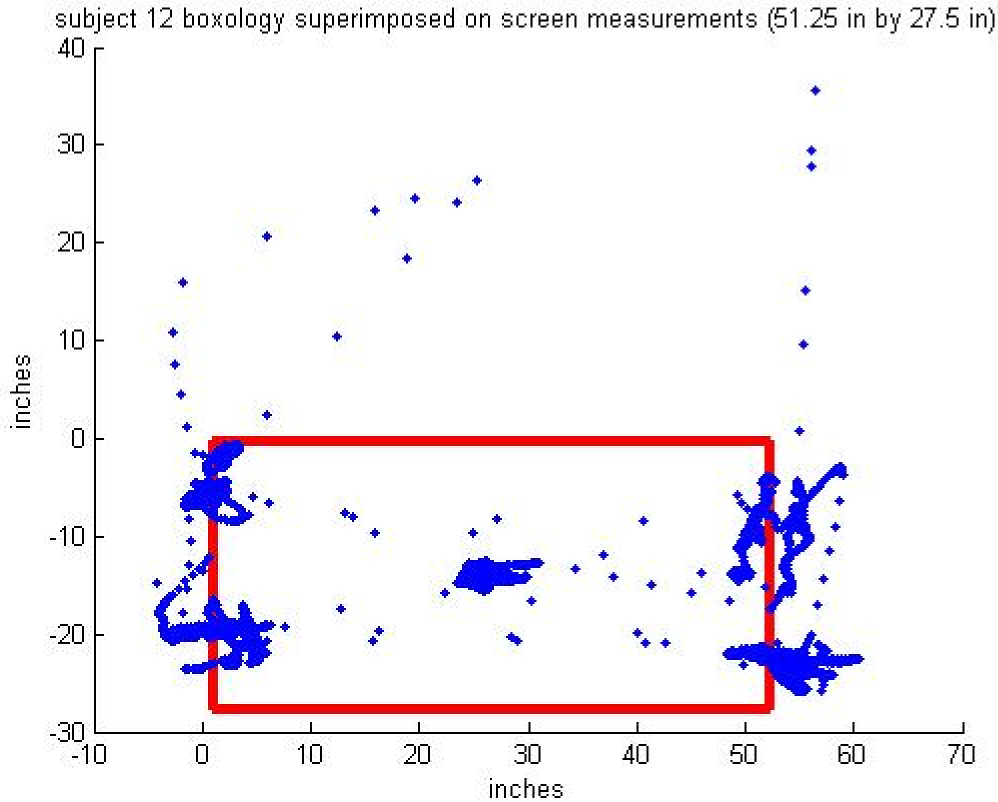


Figure 26. Boxology Data Transformed into Inches (Subject 12 Example)

A red rectangle of these dimensions is shown in Figure 26 with boxology data superimposed on top. The position of the red rectangle was found such that the center of the red rectangle coincides with the center of the middle gaze cluster of the boxology data. While the distance transformed boxology data shows reasonable agreement with the distances measured on the screen, the absolute position for the boxology tends not to be consistent from trial to trial, as mentioned above. Also, while the yaw measurement produces x-axis distances which look reasonable, the pitch measurement seems a bit too small. Generally, if one is looking for clusters within the yaw/pitch data, it is better to stick closer to the raw, non-derived Facelab measurements rather than dealing with a transform of the pitch and yaw measurements.

8.1.9 References

Seeing Machines (2005). FaceLab 4 User Manual. Canberra, Australia: Seeing Machines.

8.2 Appendix B: Speed Profile Evaluation

8.2.1 Evaluating Delay with Cross-Correlation

For the car following task, the input signal is, $f_{in}(t)$, the speed of the target vehicle, and the output signal is $f_{out}(t)$, the speed of the subject's vehicle. Figure 27 shows an actual example of these signals for subject 12 run 5. One of the measures of subject performance is an aggregate delay between $f_{in}(t)$ and $f_{out}(t)$.

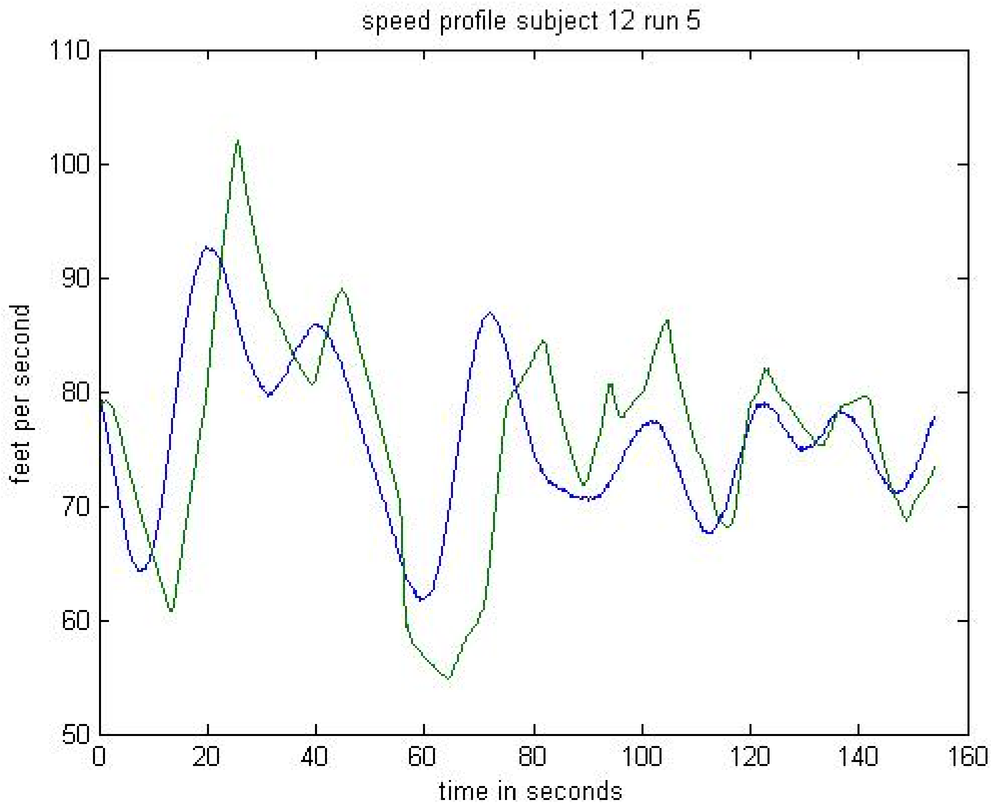


Figure 27. Sample Speed Trace of Subject Vehicle Following the Lead Vehicle

8.2.2 Cross-Correlation

This delay may be found by finding the peak of the cross-correlation between a zero mean $f_{in}(t)$ and $f_{out}(t)$. The cross correlation is expressed in continuous time as:

$$F(d) = \int_{-\infty}^{\infty} f_{in}(x) f_{out}(x+d) dx$$

In order to find the cross correlation between the above functions $f_{in}(t)$ and $f_{out}(t)$, we subtract the means from the above functions and use the Matlab function "xcorr" to obtain the plot shown in Figure 28:

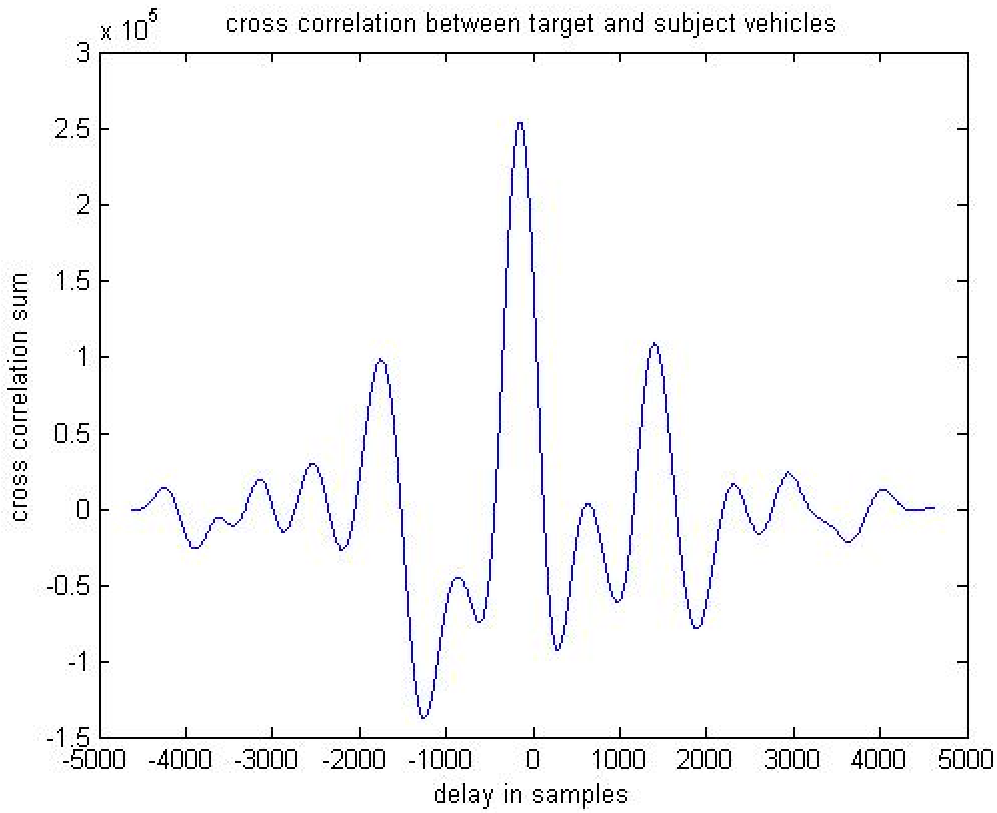


Figure 28. Cross Correlation between Subject and Lead Vehicles

8.2.3 Cross-Correlation Peak

The peak of the cross correlation function in Figure 29 occurs at an offset of -149 samples. At 30 samples per second, this is a delay of -149/30=-4.9667 seconds. Since the input functions are periodic, there are several, smaller peaks corresponding to "matches" which are integral numbers of period offsets from the maximum offset. As a sanity check, we plotted the functions in Figure 27 with a -149 sample delay.

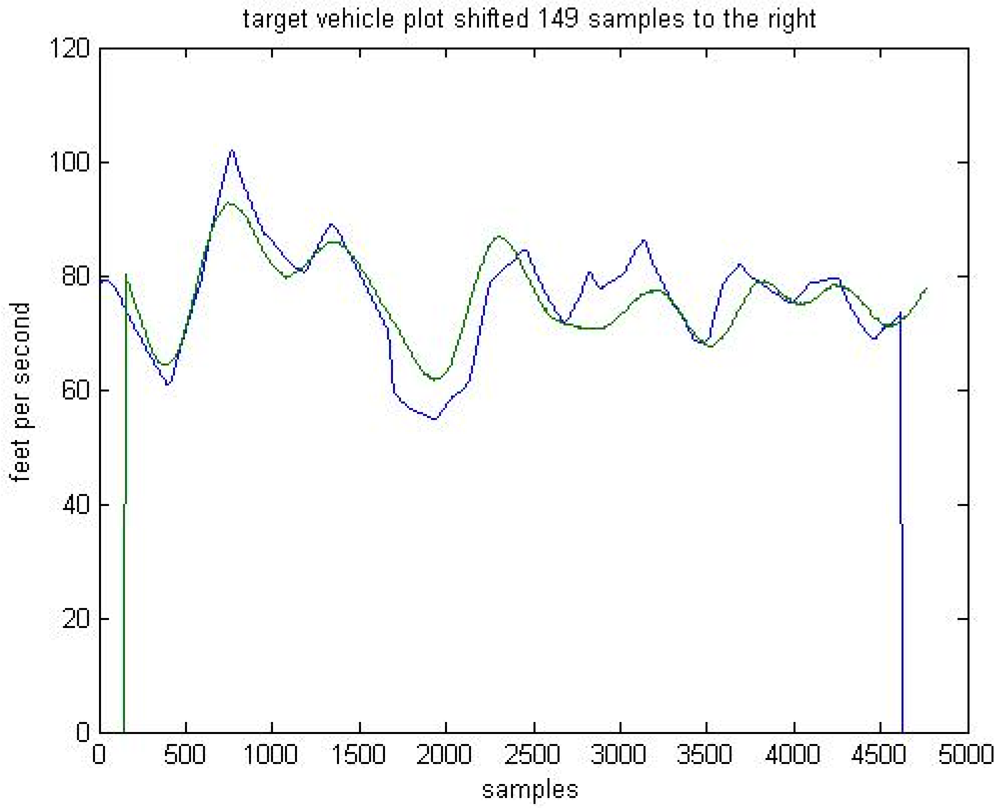


Figure 29. Peak of Cross Correlation Function (Offset of -149 Samples)

Figure 29 shows reasonable agreement between the two functions after shifting the input function 149 samples (4.9667 seconds) to the right. The nearly vertical lines on the right and left sides of the plot are artifacts of padding the plots with zeroes.

8.2.4 Cross-Correlation's Advantages over Frequency Domain Techniques

Rather than transferring $f_{in}(t)$ and $f_{out}(t)$ into the frequency domain and comparing the resulting $F_{in}(\omega)$ and $F_{out}(\omega)$, the cross-correlation technique described above allows an analyst to compare the input signal directly with the output signal without using the frequency domain as an intermediate step. The cross-correlation method does not have the frequency domain's windowing issues, nor does the delay estimation require the selection of a frequency of interest.

8.2.5 A Cross-Correlation-Based Similarity Index

Once a region of overlap has been established by finding the delay, the magnitude of the maximum cross-correlation may give some idea what degree of similarity the waveforms possess.

In discrete time, the cross-correlation at a particular delay d may be found with the dot product:

$$c_d = \vec{a} \cdot \vec{b}$$

where $\vec{a}$ and $\vec{b}$ are vectors representing the points of the overlapping portions of the sampled versions of $f_{in}(t)$ and $f_{out}(t)$ after the aggregate delay d has been found and applied. Unfortunately, the raw cross-correlation shown above is not particularly illuminating because the scale of $\vec{a}$ and $\vec{b}$ is arbitrary. For the purpose of comparing the maximum cross-correlation of two separate matches, it would be better to normalize $\vec{a}$ and $\vec{b}$ such that $\left|\vec{a} \cdot \vec{b}\right| \leq 1.0$ and $\vec{a} \cdot \vec{b} = 1.0$ for $\vec{a} = \vec{b}$.
The similarity index:

$$c_d = \frac{\vec{a}}{|\vec{a}|} \cdot \frac{\vec{b}}{|\vec{b}|}$$

will be 1.0 when $\frac{\vec{a}}{|\vec{a}|} = \frac{\vec{b}}{|\vec{b}|}$ and less than 1.0 otherwise.

8.3 Appendix C: Participant Information Summary for Simulator Protocols

The following is a copy of the Participant Information Summary document used in Experiments 1 and 3. The Participant Information Summary for Experiment 2 was similar, except that it was a page shorter due to not having to describe the simulated driving components.

PARTICIPANT INFORMED CONSENT FORM

STUDY TITLE: Development of NCAP Distraction Test

STUDY INVESTIGATOR: Thomas A. Ranney, Ph.D.

STUDY SITE: Transportation Research Center, Inc.
Data Collection Annex
8200 Business Way
Plain City, OH 43064

TELEPHONE: 800-262-8309

SPONSOR: U.S. Department of Transportation
National Highway Traffic Safety Administration

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. You are being asked to take part in this study because the study investigator feels that you meet the qualifications of the study.

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 given a signed copy of your consent 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 caused by "in-vehicle technologies." The latest in-vehicle technologies include devices that provide services such as access to the internet and navigation systems (for maps and driving directions), as well as the ability to send and receive e-mails. 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 of in-vehicle technologies while driving.

Participant's Initials________________

STUDY REQUIREMENTS

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

- You are 25 – 50 years of age,
- You have a valid, unrestricted U.S. driver's license (except for restrictions concerning corrective eyeglasses and contact lenses),
- You have a minimum of two years driving experience,
- You drive at least 7,000 miles per year, and
- You are in good general health.

NUMBER OF STUDY SITES AND STUDY PARTICIPANTS

This study will take place at one research site (Transportation Research Center Inc. Data Collection Annex) and will include at least 28 participants.

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 4 hours. During this session, you will be asked to complete specific driving objectives while performing different in-vehicle tasks. 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.

Simulated Driving:

During your session, you will be asked to drive a fixed-base simulator. A fixed-based simulator is a machine that imitates the conditions of driving in real life, but does not move. The simulator will be connected to the study vehicle, which will be a recent model-year passenger vehicle (sedan, minivan, or SUV). While driving the simulator, you will sit in the driver's seat of the study vehicle. The study vehicle will have its engine turned off. You will control the simulator by moving the steering wheel and the gas and brake pedals of the study vehicle.

The study vehicle will be 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.

Participant's Initials__________________

Driving Objectives:

While operating the simulator, you will be asked to perform specific driving tasks. These tasks may involve activities such as following a car, changing lanes, and/or detecting simple targets that appear either in your peripheral (side) vision or on the computer-generated roadway image. In one condition, the target will be "head-mounted," which will require you to wear a lightweight headband with a small target attached to it.

In-Vehicle Tasks:

While completing the driving objectives, you will be asked to perform specific in-vehicle tasks. These tasks will imitate or be similar to the actions required to operate in-vehicle technologies (such as a stereo, the internet, or a navigation system).

The in-vehicle tasks will consist either of tasks using a small computer screen located inside the study vehicle, tasks that involve listening and responding verbally, or tasks using the stereo or navigation system in the study vehicle.

Visual Occlusion plus Auditory Tracking:

You may also be asked to complete a combination of visual occlusion and auditory tracking while performing the in-vehicle tasks in the test vehicle. Visual occlusion requires that you wear a set of glasses, which have lenses that can be made to be either transparent or opaque. An electrical current can quickly change the glasses between these two states. When they are transparent, you will be able to see normally; however, when they are opaque, you will not be able to see through them. The glasses are connected to a computer, which controls when the lenses change between opaque and transparent. When you are wearing these glasses, you will also perform an auditory tracking task, which requires that you use a joystick to follow a path that you cannot see. Sounds presented in the left or right speaker will help you to follow the imaginary path. The visual occlusion and auditory tracking tasks are intended to replace the requirement to drive the simulator.

Eye Movement Recording and Monitoring:

Video cameras will be used to monitor your eye movements while operating the driving simulator and performing the in-vehicle tasks. The video cameras are located so that they will not affect your driving. The information collected using these video cameras is recorded so that it can be analyzed at a later time.

There are certain requirements for accurately recording your eye movements while driving. These requirements are as follows:

- Your entire face must be clearly visible while driving. If your hair hangs in your face, you may be asked to use clips or a rubber band to keep it out of your face.
- If you require corrective lenses and have contact lenses, you will be asked to wear them rather than glasses.

Participant's Initials_________________

- To help the eye tracking system better identify and track your facial features, you will be required to wear several small stickers on your face. The stickers will be put on before you begin driving and cannot be removed or moved until a member of the study staff informs you that you are finished driving. As a result, you may be wearing the stickers for up to 3 hours.

Summary of Study Procedures:

The following procedures will take place at your session:

- After signing this consent form, you will be given instructions, training, and practice time for driving the simulator and performing visual occlusion and in-vehicle tasks.
- You will then complete a number of short tests, each lasting approximately 3 minutes. Each test will involve a different combination of driving objectives and in-vehicle tasks. You will be asked to complete approximately 35 tests (including all tests completed during training and practice).
- At the conclusion of the tests, you will be asked to answer brief questions about the tasks that you performed.
- After completing the questions, the session will end and your participation in this research study will be complete.

NEW INFORMATION

We do not anticipate that any changes to procedures will take place during this study. However, any new information developed during the course of the research that may affect your willingness to participate will be provided to you.

RISKS

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 quit driving as a result of discomfort, you will be allowed to quit at once. You will be asked to sit and rest before leaving, while consuming a beverage and a snack. There is no evidence that driving ability is hampered in any way; therefore, if you 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.

You will be asked to wear several small latex stickers on your face while driving. These stickers may cause skin irritation in people with an allergy to latex. Allergic reaction may be mild (rash, hives) to severe (difficulty breathing, or a collapse of blood circulation and breathing systems).

Participant's Initials________________

A severe allergic reaction, which is extremely unlikely, would require immediate medical treatment and could result in permanent disability or death.

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

BENEFITS

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. Your participation in this study will provide data that may help develop these recommendations or standards.

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 PARTICIPATION, WITHDRAWAL, AND TERMINATION

Participation in this research is voluntary. By agreeing to participate, you agree to operate the research vehicle 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 withdrawn from the study. You may withdraw your consent and discontinue participation in the study at any time without penalty.

COSTS TO YOU

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

COMPENSATION

You will receive $26.00 per hour for the time you spend at the data collection facility. In addition, you will have the opportunity to earn incentive pay based on your performance on the driving and in-vehicle tasks. The maximum possible amount of incentive pay is approximately $9.00 per hour depending on the specific number of tests completed.

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

Participant's Initials__________________

USE OF INFORMATION COLLECTED

In the course of this study, the following data will be collected:

- Engineering data (such as the information recorded by the study vehicle sensors)
- Video/audio data (such as the information recorded by the video cameras)

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 in final reports or other publication or media for scientific, education, research, or outreach purposes.

The **video/audio data** recorded in this study includes your video-recorded likeness and all 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.

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.

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, you may contact Dr. Sally P. Green, Chairman of Sterling Institutional Review Board, 6300 Powers Ferry Road, Suite 600-351, Atlanta, Georgia 30339 (mailing address) at telephone number 1-888-636-1062 (toll free).

Participant's Initials_______________

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 vehicle 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 provided 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. (TRC) employees associated with this study during the pre-participation phone interview 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 vehicle in accordance with all instructions provided to me by the study staff.

I have been told that:

- The study will be conducted on 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.

Participant's Initials________________

- 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 (937) 666-4511 or 800-262-8309 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 have not waived any of my legal rights by signing this document.

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 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. NTHSA 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 Traffic
(*Printed Name of Participant*)
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

 Participant's Initials__________________

8.4 Appendix D: Rating Scale Mental Effort (RSME)

Instructions

We are interested not only in assessing your performance but also the experiences you will have during the different task conditions. Right now I will describe the technique that will be used to examine your experiences.

Most importantly, we want to assess the mental effort you experience. Mental effort is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of mental effort may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The mental effort contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another.

Since mental effort is something experienced individually by each person, there are no effective “rules” that can be used to estimate the mental effort of different activities. One way to find out about mental effort is to ask people to describe the feelings they experienced. We will be using a rating scale to assess your mental effort. Please read the definition of the scale carefully. If you have a question about the scale, please ask me about it. It is extremely important that it is clear to you. The description will be made available to you for reference during the experiment.

Rating Scale Definition
Mental Effort: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving? How hard did you have to work mentally? How much time pressure did you feel?

After performing a set of tasks, you will be instructed to bring the vehicle to a stop at a specified location. While the vehicle is stopped, the rating scale will be presented to you. You will evaluate the tasks performed (some combination of car following, light detection and phone tasks) since the time when the previous rating scale was administered, by telling the in-vehicle experimenter the number on the scale at the point that matches your experience. Please consider your responses carefully in distinguishing among the different task conditions. Your ratings will play an important role in the evaluation being conducted, thus your active participation is essential to the success of this experiment, and is greatly appreciated.

Subject: Date:
Condition: Exp:

Rating Scale Mental Effort

Modified 8/25/03

Please indicate, by marking the scale below, how much effort it took for you to perform the task you just completed

10	Extreme Effort
9	
8	Great Effort
7	
6	Moderate Effort
5	
4	Some Effort
3	
2	Little Effort
1	
0	Absolutely No Effort

Rating Scale Definition
Mental Effort: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving? How hard did you have to work mentally? How much time pressure did you feel?

8.5 Appendix E: Simulator Sickness Questionnaire

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 _____________________________________

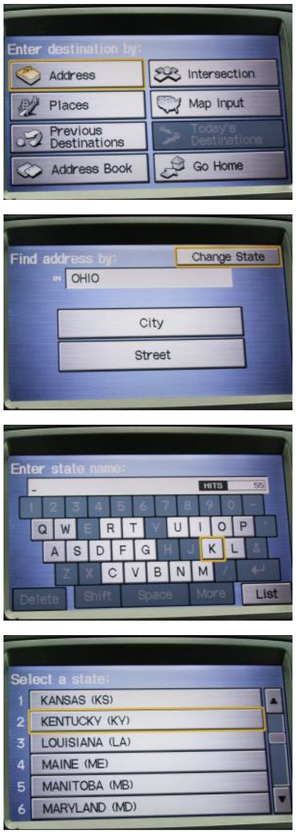
* Vertigo is experienced as loss of orientation with respect to vertical upright.

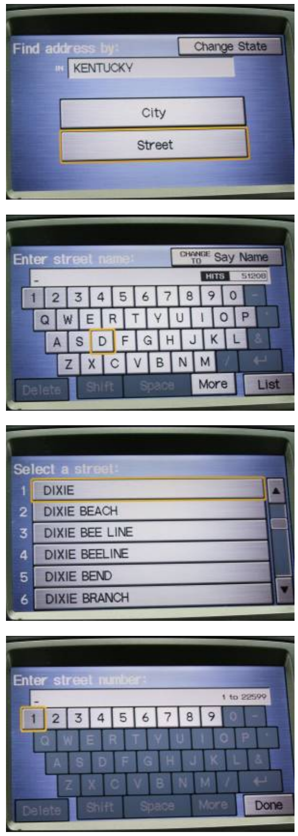
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

8.6 Appendix F: Navigation System Comparison

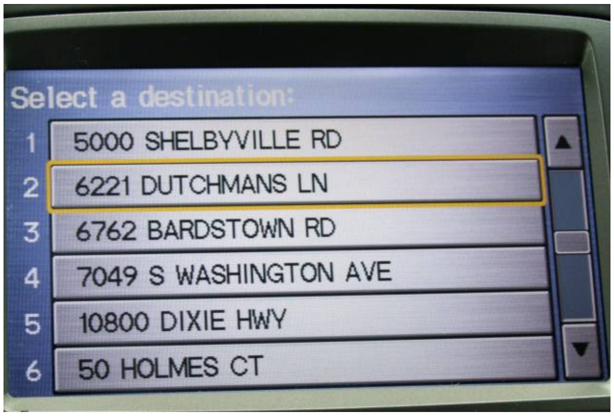
Interaction	OEM	High-Usability	Low-Usability
01) State Input	Name (w/ Autocomplete)	Name Only (Autocomplete requires many letters)	Scroll through list of states (No Autocomplete)
02) City Input	First 3 or 4 letters (Autocomplete)	First 3 or 4 letters (Autocomplete)	Entire name (No Autocomplete) list has many
03) Street Input	First 3 or 4 letters (Autocomplete)	Name (No Autocomplete)	Name (No Autocomplete) Always prompts street type?
04) Address Input	Type	Type	Type
05) Order of Input	St > City > Adr > Str	St > City > Adr > Str	St > City > Str > Adr
06) Input Completion	View destination	Map, back button to main or Go, hit the menu button after	Map, back button to main or Go, hit the menu button after
07) Remote	No (?)	Yes	Yes
08) Error Detection (for unintended key presses)	No	Yes	No
09) Previous Destinations	Organized by entry time	Organized by distance	Organized by abc or distance
10) System Response Delay	Medium SRD	Low SRD	Low SRD
11) LS, PD	LS – Separate city mode PD – Previous Destinations	LS - Separate city mode PD - Favorites	LS - Access in city center PD – Saved Places
13) Keyboard	Qwerty or Alphabetic	Qwerty or Alphabetic	Alphabetic
14) Error Correction	Pushing the cancel button located to the left of the navigation screen moves the system to the previous input type. (It is cleared).		Pushing the back button in the lower left moves the system immediately to the previous screen.

8.7 Appendix G: Display Screen Images of 3 Navigation Systems, Experiment 3



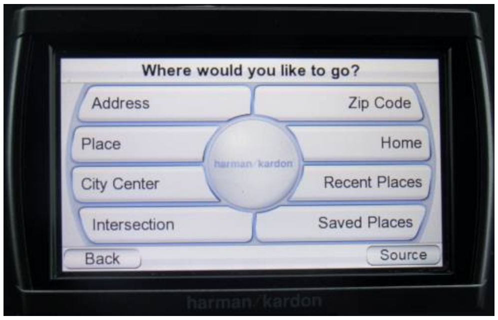


OEM System Address Entry (ADD)

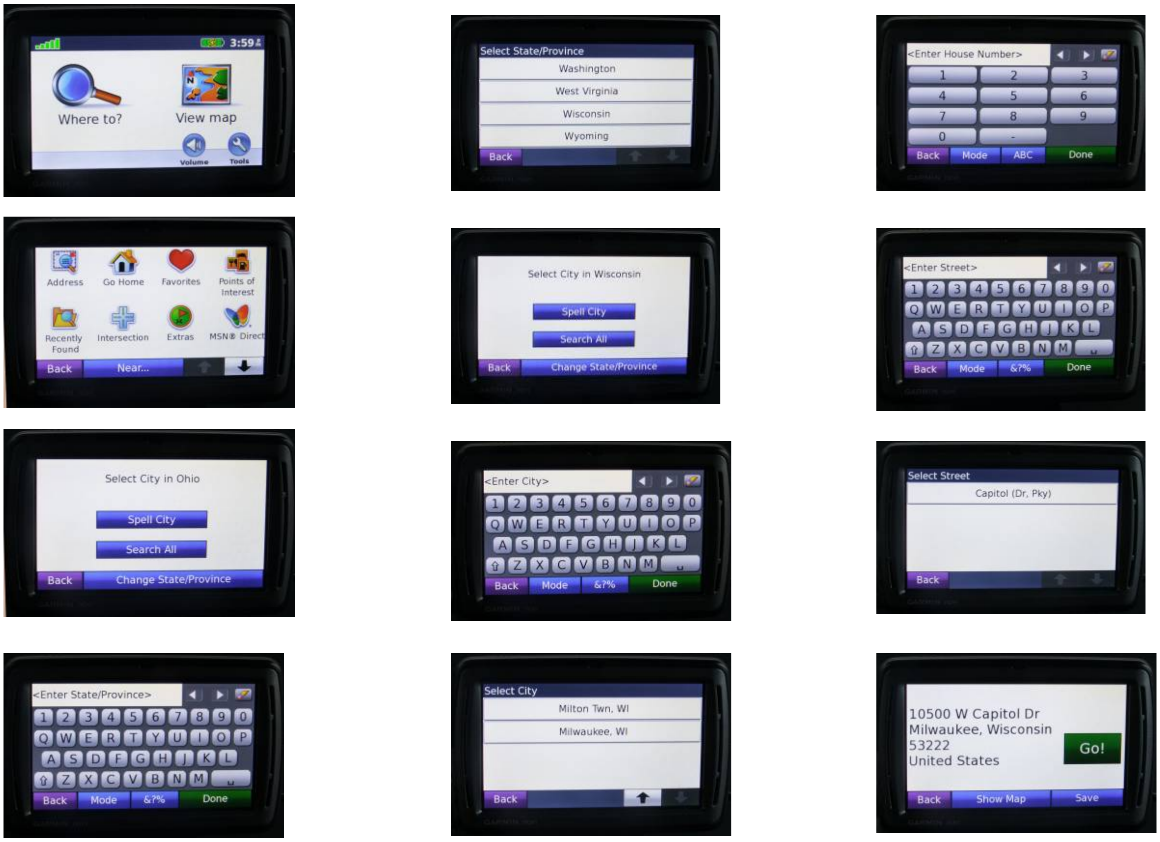


OEM System
List Search
(PD)

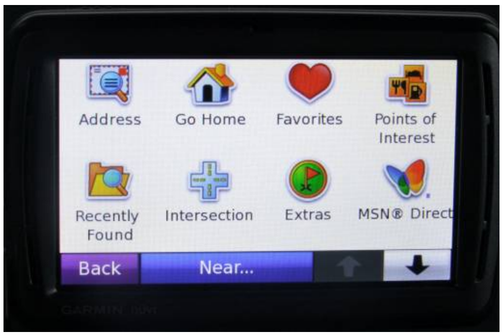
Low-Usability
Address Entry
(ADD)

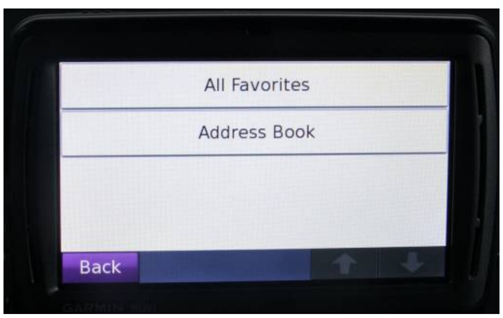


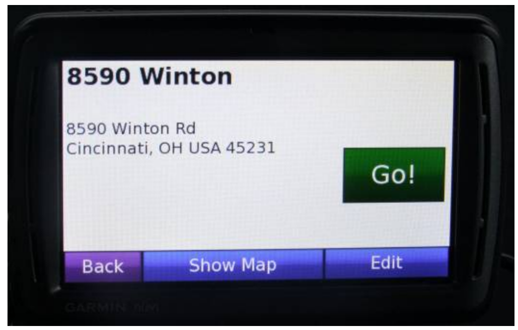
Low-Usability
List Search (PD)



High-Usability
Address Entry
(ADD)







High-Usability List Search (PD)

DOT HS 811 463
November 2011

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

NHTSA
www.nhtsa.gov