PB2000-106898

U.S. Department
of Transportation

**National Highway
Traffic Safety
Administration**

DOT HS 809-069

July, 2000

Driver Distraction with Wireless Telecommunications and Route Guidance Systems

REPRODUCED BY: NTIS
U.S. Department of Commerce
National Technical Information Service
Springfield, Virginia 22161

This publication is distributed by the U. S. Department of Transportation, National Highway Traffic Safety Administration, in the interest of information exchange. The opinions, findings, and conclusions expressed in this publication are those of the author(s) and not necessarily those of the Department of Transportation or the National Highway Traffic Safety Administration. The United States Government assumes no liability for its contents or use thereof. If trade or manufacturers' names or products are mentioned, it is because they are considered essential to the object of the publication and should not be construed as an endorsement. The United States Government does not endorse products or manufacturers.

PROTECTED UNDER INTERNATIONAL COPYRIGHT
ALL RIGHTS RESERVED
NATIONAL TECHNICAL INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE

Reproduced from best available copy.

Technical Report Documentation Page

1. Report No. DOT HS 809-069	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Driver Distraction with Route Guidance Systems		5. Report Date July, 2000
		6. Performing Organization Code
7. Author(s) L. Tijerina, S. Johnston, E. Parmer, and M. D. Winterbottom, Transportation Research Center Inc. (TRC) Mike Goodman, 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 400 Seventh Street, S.W. Washington, D.C. 20590		13. Type of Report and Period Covered Technical report
		14. Sponsoring Agency Code NHTSA/NRD-22

15. Supplementary Notes

The authors thank Frank Barickman (NHTSA), Duane Stoltzfus (TRC), and Steve Wilson (Sophisticated Systems Inc.) who developed instrumentation hardware and software needed for this research. Mark Gleckler, Heath Albrecht, and Adam Andrella provided significant technical support for vehicle instrumentation, data processing and reduction, and data management. Riley Garrott, August Burgett, Duane Perrin, and Joseph Kanianthra of NHTSA provided valuable technical guidance.

16. Abstract

Concerns have been raised in recent years about the distraction potential of Intelligent Transportation Systems (ITS) technologies including driver information systems such as route navigation systems. The research described in this report had the following objectives: 1) characterize the impact of route guidance system destination entry use on vehicle control and driver eye glance behavior on a test track; 2) assess the influence of individual differences, as indexed by a battery of cognitive tests, on the susceptibility to distraction as indicated by disruption in vehicle control and driver eye glance behavior during destination entry and cellular telephone use while driving; and 3) examine the validity of a proposed SAE recommended practice, known as the 15-second rule, according to which if a given route guidance destination entry function can be completed in 15 seconds or less by a sample of drivers without concurrent driving, then that function may be accessible while the vehicle is in motion.

Results for this research suggest voice recognition technology is a viable alternative to visual-manual destination entry while driving and that destination entry with visual-manual methods is ill-advised while driving. The assessment of the impact of individual differences on the susceptibility to distraction during destination entry and cellular telephone use while driving showed low but consistent patterns of correlation to test track performance measures. The results of this preliminary assessment of the 15-second rule suggest that, when applied to a variety of tasks, the rule has diagnostic sensitivity not much better than chance guessing. The 15-second rule works well for disallowing the most egregiously distracting tasks, e.g., manual destination entry while driving. These preliminary findings, together with the observation that the 15-second rule is, in itself, not diagnostic with regard to the locus of a driver distraction effects, suggest that opportunities for improvement should be pursued.

17. Key Words Driver Workload, Cellular Telephone, Crash Avoidance, ITS		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price

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

Table of Contents

List of Figures v
List of Tables vi
Executive Summary vii

1.0 General Introduction to Driver Distraction with Cellular Telecommunications and Route Guidance Systems 1
1.1 Background 1
1.2 Review of Previous Research: Cellular Telecommunications Systems 1
1.2.1. Laboratory Studies 1
1.2.2 Simulator and Test Track Research 2
1.2.3 On-road Studies 3
1.2.4 Epidemiological Studies 4
1.3 Review of Previous Research: Route Guidance Systems 5
1.3.1 Route Guidance System Design and Driver Workload: Destination Entry 6
1.3.2 Basic Driver Needs for Wayfinding 8
1.3.3 Studies of Route Guidance System Design and Driver Workload: Route Guidance 9
1.3.4 Route Guidance Systems Interfaces: Modes and Codes 15
1.3.5 Verbal Route Guidance as a Means to Reduce Driver Workload 16
1.3.6 Use of Landmark Information in Route Guidance Systems 19
1.4 Voice Recognition as a Means to Reduce Driver Workload 21
1.5 Status of Commercially Available Route Guidance Systems 25
1.6 Research Objectives 27
1.7 Organization of the Report 27
1.8 Summary 27

2.0 Driver Workload Assessment of Route Guidance System Destination Entry While Driving: a Test Track Study 28
2.1 Introduction 28
2.2 Approach 28
2.2.1 Test Participants 28
2.2.2 Test Vehicle and Instrumentation 28
2.2.3. Route Guidance Systems 28
2.2.4 Test Route 29
2.2.5 Independent Factors, Dependent Measures, and Study Design 29
2.2.6 Procedure 30
2.2.7 Data Analysis 30
2.3 Results 30
2.4 Conclusions 32

3.0 Individual Differences and In-vehicle Distraction While Driving: a Test Track Study and Psychometric Evaluation 35
3.1 Introduction 35

Table of Contents – Continued

3.2 Approach 35
3.2.1 Test Participants 35
3.2.2 Test Vehicle and Instrumentation 35
3.2.3. Route Guidance Systems 36
3.2.4 Test Track 38
3.2.5 Procedure 38
3.2.6 Test Track Measures, Test Battery Measures 38
3.3 Results 39
3.4 Discussion 41

4.0 Preliminary Evaluation of the Proposed SAE J2364 15-second Rule for Accessibility of Route Navigation System Functions While Driving 42
4.1 Introduction 42
4.2 Approach 43
4.2.1 Test Participants 43
4.2.2 Test Vehicles and Test Track 43
4.2.3 Route Guidance Systems and Other In-Vehicle Tasks 44
4.2.4 Response Measures 44
4.2.5 Procedures 45
4.2.6 Data Analysis 45
4.3 Results 45
4.3.1 The Relationship Between Static Completion Time and Dynamic Completion Time 46
4.3.2 The Relationship between Static Completion Time and Lane Exceedences 47
4.3.3 The Relationship between Dynamic Completion Time and Lane Exceedences 47
4.3.4 Signal Detection Analysis of the 15-Second Rule 47
4.4 Discussion 55
4.4.1 On the Nature of Completion Times 55
4.4.2 Lanekeeping Performance and Completion Times 55
4.4.3 Classification Accuracy of the 15-Second Rule 56
4.5 Recommendations 56
4.5.1 Necessary and Sufficient Safety-Relevant Measures of Driving Performance 57
4.5.2 Beware the Chain of Causal Inference 58
4.5.3 Eye Glance Measurement is Not Necessarily Too Difficult 58
4.5.4 Efficient Driver Performance Measurement is Feasible 58
4.5.5 Static Completion Time May be Misleading as a Link to GOMS Modeling 59
4.5.6 The Imperfect Evaluation Rule is Better Than Nothing At All... 59
4.6 Postscript 63

Table of Contents – Continued

5.0 References 64

Appendix A 75

Appendix B 87

List of Figures

Figure 2.1. Age and Device Effects on Trial Time (i.e., Task Completion Time) 33
Figure 2.2. Device and Note Card Effects on Glance Frequency and Duration 33
Figure 2.3. Age and Device Effects on Number of Lane Exceedences per Trial 34
Figure 2.4. Age and Device Effects on Eyes-Off-Road-Ahead Time and Road Glance Duration 34
Figure 4.1. Regression and Scatter Plot of Dynamic Completion Time vs. Static Completion Time 46
Figure 4.2. Regression and Scatter Plot of Lane Exceedences per Trial vs. Static Completion Time 47
Figure 4.3. Regression and Scatter Plot of Number of Lane Exceedences vs. Dynamic Completion Time 48
Figure 4.4. Static Completion Time ROC Curve 52
Figure 4.5. Scatter Plot of Static Completion Time for Each Device 54

List of Tables

Table 1-1. Navigation System Features for Systems Reviewed in Paul (1996) 26
Table 3-1. PATSYS Test Battery: Temporal and Cognitive Subtests 37
Table 3-2. Intercorrelation Matrix for Test Track and Test Battery Measures. 40
Table 4-1. Calculation of the Area Under the ROC Curve for the Static Completion Time Measure 51
Table 4-2. Classification Results of 15-Second Rule 53

EXECUTIVE SUMMARY

Recent trends in technology suggest that the future of in-vehicle "telematics" will involve products which fully integrate telecommunications, route guidance, and other driver information systems functions, perhaps with a voice recognition-based driver interface. The studies and analyses described in this report represent the authors' attempts to contribute to a better understanding of the complex relationship between the driver and the technology. It is hoped that this understanding, in turn, supports a more driver-centered evolution of the technology.

Clearly, a great many issues arise in consideration of the human factors of wireless telecommunications and route guidance systems for drivers. The following objectives were selected for the studies described in this report:

- Characterize the impact of route guidance system destination entry on vehicle control and driver eye glance behavior on a test track;
- Assess the influence of individual differences, as indexed by a battery of cognitive tests, on the susceptibility to distraction as indicated by disruption in vehicle control and driver eye glance behavior during destination entry and cellular telephone use while driving;
- Examine the validity of a proposed SAE recommended practice to assess whether or not a given route guidance destination entry function ought to be allowed while the vehicle is in motion;

A test track study was performed which examined four commercially available route guidance systems, representing alternative destination entry and retrieval methods, in terms of driver visual allocation, driver-vehicle performance, and driver subjective assessments. Results for this study suggest voice recognition technology is a viable alternative to visual-manual destination entry while driving. This result is highlighted in test participant subjective assessments that favored voice input over visual-manual methods. These data suggest that destination entry with visual-manual methods is ill-advised while driving.

The influence of individual differences on driver distraction was examined on a test track. Subjects were trained on destination entry procedures with four commercially available route guidance systems, as well as the dialing task on a commercially available wireless cellular telephone and on manually tuning an after-market car radio. The participants then drove an instrumented vehicle on a test track while concurrently engaging in various tasks with these devices. In-vehicle task completion time, average glance duration away from the road ahead, number of glances away from the road ahead, and number of lane exceedences were recorded. The participants were later given an automated battery of temporal visual perception and cognitive tasks. Performance on the test battery was then correlated to performance on the test track measures to determine the extent to which individual driver differences could account for observed performance differences. Analysis of these elementary test scores as predictors show low but consistent patterns of correlation to test-track performance measures.

The draft standard SAE J2364 proposes that if a specified route-guidance destination entry function can be completed by a sample of drivers in 15 seconds or less without concurrent driving, then that

function may be accessible to a driver in a moving vehicle. To assess the diagnostic properties of this so-called 15-second rule, another was conducted. Ten subjects between the ages of 55 and 69 years completed 15 different tasks in a stationary vehicle (15-second rule) and while driving. Tasks included various destination entry tasks with four commercially available route guidance systems, manual cell phone dialing, manually tuning an after-market in-dash radio to specific AM and FM stations, and adjusting the HVAC controls in the test vehicle. Results were characterized in terms of the diagnostic sensitivity of the 15-second rule. There were four possible outcomes: (1) True Positives, in which a task both failed the 15-second rule and showed an adverse effect on driving performance; (2) False Negatives, in which a task passed the 15-second rule but showed an adverse effect on driving performance; (3) True Negatives, in which a task passed the 15-second rule and showed no adverse effects on driving performance, and (4) False Positives, in which a task failed the 15-second static test but did not adversely affect driving performance. Disrupted lanekeeping refers to an increased likelihood of exceeding the lane boundaries during performance of the task in a moving vehicle. The main results are summarized below:

(1) True Positives: All route navigation system destination entry tasks that required visual-manual methods both failed the 15-second rule and were associated with disrupted lanekeeping. Manually dialing an unfamiliar 10-digit phone number on the cellular telephone also fell into this classification.

(2) False Negatives: Tuning the Clarion after-market radio took less than 15 seconds for static completion yet was associated with above-threshold disruptions of lanekeeping on the test track.

(3) True Negatives: The HVAC adjustment was the only task that was both completed in less than 15 seconds statically and had no appreciable effect on lanekeeping during the test track trials.

(4) False Positives: Dialing home on a cellular telephone, all VAAN destination entries by voice, and tuning the after-market radio to a prescribed FM station all took longer than 15 seconds to complete in a parked vehicle (i.e., statically) yet were not associated with significant disruptions in lanekeeping on the test track.

The estimated area under the receiver operating characteristic (ROC) curve was 0.55, indicating close to chance classification sensitivity. Thus, the results of this preliminary assessment suggest that when applied to a variety of in-vehicle tasks, the 15-second rule has diagnostic sensitivity not much better than chance guessing. It produces a high false positive rate while still allowing false negatives. It should be further noted that the 15-second rule has not yet been validated with respect to object and event detection, the effects of which can be distinct from degraded vehicle control. The 15-second rule works well for disallowing the most egregiously distracting tasks, e.g., manual destination entry while driving, for which the rule was intended. However, with few exceptions, a 30-Second or a 45-Second Rule would also work about as well. These preliminary findings, together with the observation that the 15-second rule is, in itself, not diagnostic with regard to the locus of driver distraction effects, suggest that opportunities for improvement or enhancement should be pursued. Recommendations are discussed.

1.0 General Introduction to Driver Distraction with Cellular Telecommunications and Route Guidance Systems

1.1 Background

Concerns have been raised in recent years about the distraction potential of Intelligent Transportation Systems (ITS) technologies. Such technologies include wireless telecommunications devices such as cellular telephones and driver information systems such as route navigation systems. This concern prompted the National Highway Traffic Safety Administration (NHTSA) to recently release a major report on the safety implications of the use of cellular telephones while driving (Goodman, Bents, Tijerina, Wierwille, Lerner, and Benel, 1997; Goodman, Tijerina, Bents, and Wierwille, 1999). Similarly, numerous research programs have been carried out by government, academia, and industry to investigate human factors and safety issues associated with route guidance systems. Most recently, voice recognition systems have been offered for use in cars and trucks in the hope that such technology will largely alleviate driver distraction. A synopsis of selected human factors studies into cellular telephone and route guidance system use is provided below. This is followed by a discussion of key issues in driver distraction associated with route guidance and voice recognition systems. This review will serve as a general introduction to the area of device-induced driver distraction. This section will conclude with the research objectives selected for the series of studies described in this report and a description of the organization of the report.

1.2 Review of Previous Research: Cellular Telecommunications Systems

The extensive growth in the wireless communications industry over the past ten years has been accompanied by growing concern for the potential hazards of drivers using wireless communication devices from moving vehicles. In response to this concern, the NHTSA initiated a comprehensive review of existing data on the topic (Goodman, Bents, Tijerina, Wierwille, Lerner, and Benel, 1997). This section provides a summary of findings of the published human factors literature on cellular telephones and driving reviewed in that report and indicates areas of uncertainty and future research needs. See Goodman, et al. (1997) for additional details.

Over two dozen studies in English were reviewed. They included laboratory studies, simulator studies, closed course/test track studies, studies conducted on open public roads, and epidemiological investigations. They spanned a range of cellular telephone tasks, from manual dialing, to various types of voice communications and comparison conditions. As will be indicated below, the methods of investigation and the materials used can have a substantial impact on the results obtained.

1.2.1 Laboratory Studies: Laboratory tasks are most problematic to interpret. For example, Boase, Hannigan, and Porter (1988) used a computer game of "squash" as a surrogate for automobile driving while studying the impact of simple dialogues (e.g., asking about favorite foods or past education) versus difficult negotiation dialogues (e.g., returning faulty goods, or responding to an altered holiday travel booking). The relevance to driving of the finding that primary task performance (measured in percent balls missed) deteriorates when conversation is introduced is hard to determine without any rigorous link being made between the computer game and driving.

Briem and Hedman (1995) used a pursuit tracking task on a computer screen as a surrogate for driving to study the effects of conversing on a cellular telephone. The authors reported a complex set of results, one part of which indicated that conversation can deteriorate tracking performance, especially when the tracking task has been made more difficult to simulate driving on slippery roads. Again, no data are provided to support the assertion that the pursuit tracking task adequately simulated driving on real roads.

1.2.2 Simulator and Test Track Research: The simulator and test-track studies deal with many facets of driver behavior and performance while using cellular telephones. With respect to the dialing task, the studies reviewed suggest the following. When compared to driving alone, cellular telephone manual dialing can be disruptive of vehicle control activities like lanekeeping and speed maintenance (Stein, Parseghian, and Allen, 1987; Zwahlen, Adams, and Schwarz, 1988; Serafin, Wen, Paelke, and Green, 1993a, 1993b). However, this disruption does not always appear, especially in closed-course environments (Kames, 1978). Studies using voice dialing emulations do generally support the hands-free approach as a desirable design goal. Manual dialing is sometimes, but not always, found to be more disruptive than manually tuning a radio (McKnight and McKnight, 1991, 1993; Stein, et al., 1987). Subjective assessments by test participants indicate that they are generally aware of the demanding nature of manually dialing a cellular telephone. Many studies report driver behavior that appear to be attempts to compensate for such disruptive effects (e.g., by slowing down the vehicle).

For the voice communications task and its effects on driving the following can be concluded. On the positive side, voice communications, if sufficiently frequent and simple to perform, appear to enhance driving performance with fatigued drivers (Drory, 1985). Equally important, simple conversations appear to have little impact on lanekeeping and speed maintenance but sometimes affect driver situational awareness (e.g., increased reaction times, reduced mirror sampling).

As a rule, however, the simulator and test track studies that make use of cognitively demanding "intelligence test" conversational materials generally show degradations in lanekeeping, speed maintenance, or headway maintenance (Alm and Nilsson, 1990; Nilsson and Alm; 1991; Nilsson, 1993; Serafin, et al., 1993). The impact of such voice communications on perceptual and judgement performance and object and event detection is also negative (e.g., Brown, Tickner, and Simmonds, 1969; McKnight and McKnight, 1991, 1993).

The correspondence between the conversational materials used in these studies and the content of normal cellular telephone communications is unknown. Thus, such results may represent worst-case or atypical voice communications. On the other hand, all simulator and test track studies to date have used conversational materials devoid of emotional content. Emotionally laden communication (e..g, a broker's call to learn that a great deal of money has been lost) may have a deleterious impact on highway safety that is even greater than that found with cognitively demanding tasks.

A better understanding of the nature of actual cellular telephone communications in business and private calls is sorely needed. This characterization would include such factors as call frequency (to both place and receive calls), call duration, call content, and call etiquette. A metric of conversational "difficulty" would also be beneficial, though a fully defensible metric may be as elusive as metrics of reading difficulty have proven to be.

1.2.3 On-road Studies: Studies conducted on open roads in normal traffic represent a significant step toward realism in testing. Based on the results of the on-road studies of cellular telephone use conducted to date, the following patterns arise. First, on the road, disruptions by manual dialing to lanekeeping or speed maintenance, as compared to manual radio tuning, appear to be small to nonexistent (Hayes, Kurokawa, and Wierwille,1989; Green, Hoekstra, and Williams, 1993; Tijerina, Kiger, Rockwell, and Tornow, 1995a, 1995b).

On the other hand, data indicate that both manual radio tuning and manual dialing can be disruptive to driving (Tijerina et al., 1995a, 1995b), and crash data indicate radio tuning is associated with crash involvement (Wierwille and Tijerina, 1994). The magnitude of demands on visual attention while dialing is sometimes less than that associated with manual radio tuning (Hayes, et al., 1989); while at other times, dialing may require a greater number of glances and total time that the eyes are off the road (Tijerina, et al., 1995a, 1995b).

Driver situational awareness (as supported by mirror sampling) appears to be reduced (Brookhuis et al., 1991; Tijerina et al., 1995), though some experimental evidence exists that this reduction occurs only under conditions where drivers judge it to be acceptable, i.e., quiet motor ways (Brookhuis et al., 1991). Cognitively demanding voice communications also appear to increase driver brake reaction times, again indicating a reduction in situation awareness.

Other related findings suggest that cellular telephone conversation need not be any more demanding than conversation with a passenger (Fairclough, et al., 1991), at least in terms of driver eye glance behavior. On the other hand, cellular telephone conversations can sometimes be more demanding than passenger conversations because the passenger can more easily adjust the nature and pace of conversation based on the current driving situation (Parkes, 1991). In addition, a driver has greater control over an in-vehicle conversation and may be reluctant to abruptly terminate a call on a cellular telephone when driving demands are high for reasons of mis-guided etiquette.

If there is a single common threat to the validity of simulator studies, it is the demand characteristics of those environments when compared to real world driving. There is currently no way to determine how closely behavior in the simulator would match behavior exhibited on the roadway other than to compare the two sets of results obtained with identical test materials and protocols. Test track studies provide a greater degree of realism than simulators but these too may be biased. Validation studies are badly needed to address this question.

One comparison of on-road study results with those obtained in a part-task simulator using the same dialing and voice communications tasks and materials led to somewhat different results (Hanowski, Kantowitz, and Tijerina, 1995; Kantowitz, Hanowski, and Tijerina, 1996). In general, it appears that in those studies, professional heavy vehicle drivers allowed the driving task to deteriorate more in the simulator than they did on the road. This suggests that the consequences of primary driving task failure on the road provide an incentive to the drivers to maintain consistent acceptable performance while driving on public roads. This incentive can be difficult to adequately emulate in the simulator environment where there are no comparable consequences to poor performance (Weimer, 1995).

It appears, then, that manual dialing can be disruptive of vehicle control performance, situational awareness and judgement. The incidence and magnitude of disruption while driving on public roads

generally appear to be less than that encountered in driving simulators or on test tracks, but may nonetheless pose a safety concern. Therefore, designs to streamline the visual-manual demand associated with cellular telephone dialing activities appear warranted.

Most of the research reviewed has been on fixed mount cellular telephones as opposed to hand-held units. Given that the majority of cellular telephones sold today are of the hand-held variety, it is not clear how generalizable the results are to the hand-helds. For example, fixed mount systems may require considerably more glance time for dialing since the driver must look away from the roadway while accessing the phone keypad. The hand-held allows the drivers to maintain the phone in a position where the roadway can be more easily monitored. However, the hand-held may require two hands to dial, in which case steering control may be compromised. Hand-helds may be stored in glove compartments, briefcases, pockets, etc., and may thus require the driver to reach and/or search for the phone. Finally, a hand-held may require that an antenna be extended, and in the case of the "flip phone," may require additional manipulation. Some 94% of cellular telephones purchased in Japan in 1995 were hand-held while the largest contributor to cellular telephone related crashes in Japan (42%) was associated with responding to a call.

Solutions to some of these concerns have focused on hands free dialing and conversation. However, the conclusions to be drawn from assessments of the effects of hands-free voice communications tasks are less clear. On-road studies indicate that if the voice communications activities have any effects at all, they are on driver situational awareness and not on vehicle control performance per se.

The simulator studies that show vehicle control disruption may reflect an experimental artifact, i.e., that drivers do not place as high a priority on the driving task in a simulator as they do on the road. The voice communications dialogue materials that have been used in this line of research often involve "intelligence test" type materials that may represent both extreme and atypical cognitive loads when compared to normal cellular telephone communications. On the other hand, all of these studies used voice communications that were free of emotional content as well. Dialogues that involve substantial degrees of conflict, for example, may be even more disruptive than the cognitively challenging materials typically included in human factors testing.

1.2.4 Epidemiological Studies: The epidemiological method can serve as a useful complement to experimental or observational research although there are certain limitations which must be kept in mind when reviewing study results. For instance, a statistical association does not necessarily imply a causal relationship and other factors such as driver personality or demographic characteristics may correlate with cellular telephone use.

Violanti and Marshall (1996) examined the amount of time per month spent talking on the phone along with eighteen other driver inattention factors. Data was obtained from 100 randomly selected drivers involved in crashes within a two year period and compared to that of a comparable control group not involved in crashes. The authors reported that talking more than 50 minutes per month on the cellular telephone while driving was associated with a 5.59-fold increase in crash risk. Crash involved subjects tended to be younger, with less driving experience and more previous crashes than non-crash involved subjects, talked longer and appeared to be engaged in more intense business calls that non-crash involved drivers.

The epidemiological method addresses the general outcome (crashes), but tells us little about why that outcome occurred. If phone use was affecting driver performance, it is not known what aspect of phone use (dialing, talking, etc.) caused the problem, and what aspect of driving performance (lane control, hazard detection, etc.) was degraded and resulted in the collision. As Violanti and Marshall (1996) were careful to point out, their methods do not even establish whether a cellular telephone was in use at the time of the collision.

A more recent epidemiological study on the relationship between cellular telephone use and traffic safety is that of Redelmeier and Tibshirani (1997). They studied 699 Toronto drivers who had cellular telephones and who were involved in motor vehicle collisions resulting in substantial property damage but no personal injury. Each person's cellular telephone calls on the day of the collision and during the previous week were analyzed through detailed billing records. The time of each collision was estimated from each subject's statement, police records, and telephone listings made to emergency services. The authors estimated that the risk of a collision was between 3.0 and 6.5 times as high within 10 minutes after a cellular-phone call began as when the telephone was not used. In addition, Redelmeier and Tibshirani reported that cellular telephone units that allowed hands-free operation offered no safety advantage.

This study is suggestive of a relationship between cellular telephone use and crashes that merits further experimental inquiry, but limitations of self-selection of study participants, variability in driving conditions and driving behavior, and no indication that the cellular telephone users were 'at fault', all limit the definitiveness of the study conclusions. The relationship between cellular telephone use and crashes is made more uncertain in light of the fact that the driver with a cellular telephone may not have been using it at the time of the crash, and where the time of the crash is itself estimated and subject to various sources of error. While Redelmeier and Tibshirani's study involves a number of shortcomings, it nonetheless represents a unique and suggestive investigation of the relationship between cellular telephone use and highway safety.

There is a great need to better understand the characteristics of cellular telephone communications (frequency, duration, content) that normally arise in the real world in order to better understand how best to represent them in future studies. It is nonetheless clear that manipulation of the cellular telephone can be disruptive of driving. It is equally clear though perhaps more insidious that voice communications can be distracting even through the driver's "eyes are on the road and hands are on the wheel." Thus, there also appears to be a need to develop better means to maintain or enhance driver situational awareness while driving. This may be accomplished through intelligent transportation system (ITS) technologies such as the "intelligent answerphone" (Parkes, 1991), driver status monitoring (drowsy driver) or other crash avoidance systems (CAS) that warn the inattentive driver of crash hazards

1.3 Review of Previous Research: Route Guidance Systems

Numerous studies have been conducted in recent years on the form and function that automotive route guidance systems should possess and on driver behavior associated with such systems. Research methods in such studies have ranged from laboratory studies on desktop PCs, to part-task driving simulator research, to instrumented vehicle studies on the open highway. Much of this

research is covered in recent literature reviews carried out for the U.S. Department of Transportation (e.g., Dingus, Hulse, Jahns, Alves-Foss, Confer, Rice, Roberts, Hanowski, and Sorenson, 1996). The research on route guidance system attributes has also recently been condensed into a number of human factors guidelines documents, most notably those of Green, Levison, Paelke, and Serafin (1995) and of Campbell, Carney, and Kantowitz (1997). Finally, an effort is currently under way to translate the existing corpus of knowledge into an industry consensus standard for route guidance systems that addresses both route guidance system design characteristics and what system functions should or should not be accessible while the vehicle is in motion (Scott, 1997). This standards development effort is being carried out under the auspices of the Society of Automotive Engineers (SAE) Safety and Human Factors Committee.

Only selected research findings are reviewed here. They serve to orient the reader to a selected set of issues that are considered by the authors to be elemental to the evolution of safe in-vehicle information and telecommunications systems. Only a few of the issues raised in this literature review have been subsequently addressed in the studies reported later. This reflects an attempt to be responsive to changing research needs as they arose during the course of this project work. The authors nonetheless believe that unaddressed issues identified and discussed in this review remain important and highly recommend them for future research.

1.3.1 Route Guidance Systems Design and Driver Workload: Destination Entry: An effort is currently being pursued by the Society of Automotive Engineers (SAE) Safety and Human Factors Committee to develop an industry consensus standard or recommended practice for route guidance systems that addresses what system functions should or should not be accessible while the vehicle is in motion (Scott, 1997). Chief among those functions of concern are destination entry functions that might be engaged while driving. Clearly, inputting a destination is the first step for obtaining benefit from a route guidance system.

It is the opinion of many in the human factors community and elsewhere that destination entry while driving is simply too distracting to be done safely. However, Green (1997) has pointed out that destination entry or retrieval en route might be attempted in any of a variety of real world scenarios, e.g.,

- the driver was in a hurry, knew the general direction in which to start, and added the destination information later;
- the driver needed to change destinations en route;
- the route guidance system does not use congestion information and a radio announcement indicates the current route is problematic;
- the driver entered the wrong destination and does not wish to stop the vehicle; or
- the driver does not know the exact destination at the beginning of the trip, enters an interim destination known to be close by, then enters the actual destination at a later time.

Thus, destination entry or retrieval while in transit may be a strong temptation to some drivers. At least one route guidance system (by Sony) includes a motion sensor that locks out such functions but this is the exception rather than the rule. Many commercially available systems do not lock out such functions while the vehicle is moving.

Driver workload has been shown to vary as a function of automotive control location and data entry method. For example, Mourant, Herman, and Moussa-Hamouda (1980) reported on the use of direct looks to in-vehicle controls of different configurations and locations as a measure of driver workload. This paper explicitly posited that: "The positioning of controls so as to minimize direct looks will permit the driver to spend more time monitoring the forward scene for potentially dangerous events." (p. 417). Mourant et al. (1980) found that the frequency of driver direct looks increased with increased hand travel distance to reach a control and also that look durations increased with increasingly complex control configurations.

Paelke (1993) used a driving simulator to examine four different touchscreen configurations for entering destination information into a simulated route guidance system: a two-button sequence for each alphanumeric entry, a QWERTY keypad layout, a phone-based keypad layout, and scrolling through an alphabetical list. Results indicated that the QWERTY and phone pad layouts were quickest to use. However, lanekeeping was significantly degraded during destination entry as compared to baseline driving. Furthermore, age effects were present such that older drivers had particular difficulty driving while entering a destination with the smaller keys and relatively greater clutter of a the QWERTY keypad layout. Monty (1984) had found the use of touchscreen keys while driving required more visual attention (longer mean glance durations) and resulted in greater driving and system task errors than conventional "hard" keys. Zwahlen, Adams, and DeBald (1988) evaluated touchscreen data entry on a closed course and found that longer eyes-off-road time were significantly correlated to greater lane deviations.

Foley and Hudak (1996) reported on an autonomous route guidance system field test for a system that accomplished destination and command entry by means of a remote keyboard unit (RKU) similar in complexity for a VCR remote control. The mean destination entry time was one minute with this device and varied from a low of 8 seconds (for destination retrieval, actually) to a high of 6 minutes and 51 seconds. The destination entry task was performed only while the vehicle was in PARK, so no data were collected on driving impacts.

An interesting alternative for data entry is voice input while the vehicle is in motion. No studies of this method of data entry while the vehicle is in motion appear to have been reported to date. However, there may be similarities between this transaction and cellular phone dialogue. At least one commercially available route guidance system offers voice-only destination entry but it requires a dialogue between the driver and the system. Tijerina, Kiger, Rockwell, and Tornow (1995) found, in their on-the-road workload study of heavy vehicle drivers using cellular telephones, that the voice communications or dialogue portion of the call can reduce mirror sampling, thus degrading situational awareness to some extent. Such findings might also arise in the case of voice destination entry as well.

An on-the-road study of driver workload during destination entry or retrieval while driving would provide an empirical assessment to substantiate a standardization effort. At the present time, the prospects for collecting data on the workload imposed by destination entry or retrieval while driving on public roads are poor. The authors recently had the opportunity to work with several commercially available route guidance systems while driving. The systems varied in their method of data or command entry. Some used bezel keys, some used remote control devices, some used rotary knobs. The systems varied considerably in their data and command entry logic: some used

alphanumeric key entry while others provided scrolling lists of destinations, for example. However, what did not vary was the impression that destination entry while driving is simply too risky for the novice system user to attempt on the open roadway. Nonetheless, some people in situations like those described by Green (1997) will attempt such transactions. Thus, characterizing the relative effects of different destination entry or retrieval methods would be beneficial. Test track testing of destination entry was chosen as a useful step in characterizing the demands associated with such activities when pursued while driving.

1.3.2 Basic Driver Information Needs For Wayfinding: Once a traveler is en route to a destination, he or she has entered into the phase of travel sometimes referred to as way-finding. Streeter, Vitello, and Wonsiewicz (1985) reported that a driver must remember three facts from any route guidance instruction: direction of turn, distance to that turn, and street to turn onto. "Notable" landmarks were also identified as useful since people indicate they rely heavily on landmarks for navigation. The turn direction information is critical since it is the basic decision the driver must make at the junction. Turn-distance information allows the driver to manage his or her attention to street signs and landmarks as the host vehicle approaches the turn area. It thus provides the driver with information to prepare to get into the appropriate lane and slow down or stop as necessary to make the turn safely and efficiently. The street name to turn onto is important so that the driver knows which of several possibly closely spaced cross streets or alleys is the appropriate one. Finally, landmark information (other than street name) can be beneficial in identifying the appropriate turning point. Waller and Green (1997) have also identified the road being driven and the approximate angle of turn as being among the basic elements of a turn display. The relative value of these information items under different driving conditions and presentations is not fully understood.

However reasonable the above set of route guidance information may seem, there is considerable variation in what is actually presented by route guidance systems. Perhaps most notably, landmark information is absent in most commercially available route guidance systems (Paul, 1996). Some systems auditory (voice) displays do not announce the street name to turn onto; others present only turn-arrows at appropriate junctions. There are differences in the sequencing of information items and the timing of their presentation as well. Thus, there are wide ranging differences in the content as well as form of route guidance instructions.

The relevance of route guidance information elements for driver workload rests in the ease with which the driver must match the information provided to the driving scene. The benefits of different information items and combinations of items presented by a route guidance system may be manifested in several measures of driver workload. In terms of driver eye glance behavior, for instance, the presence of a notable landmark may influence the amount of time drivers spend looking at a route guidance system visual display, at street signs, at the road scene, and at mirrors. Speed variability may be reduced as drivers rely less on hard-to-see street signs. Similarly, lanekeeping, car-following distance or closing rate, and lane changes may all be kept within relatively safer limits if route guidance information is optimized. At a more global level, travel time and number of wrong turns may be reduced with enhanced route guidance information. Finally, the presence of landmark information may enhance driver acceptance of a system, thereby increasing its frequency of use. All such effects have safety implications (see Tijerina, Kiger, Wierwille, and Rockwell, 1995).

1.3.3 Studies of Route Guidance System Design and Driver Workload: Route Guidance:
Electronic route guidance systems have been under development at least since the electronic route guidance system (ERGS) project pursued by the Bureau of Public Roads in the late 1960s (Rosen, Mammano, and Favout, 1970). Unfortunately, the ERGS was never field tested. Thus, it was not until some years later than human factors studies of electronic route navigation systems for surface transportation were initially conducted. A review of selected studies is provided below. See Dingus, Hulse, Jahns, Alves-Foss, Confer, Rice, Roberts, Hanowski, and Sorenson (1996) for a broader review of the human factors literature on route guidance systems.

In an early study on the subject, Streeter, Vitello, and Wonsiewicz (1985) evaluated four different route guidance aids on driver behavior and performance in an on-the-road study: voice guidance (provided through taped instructions), customized paper route maps, voice guidance plus customized paper route maps, and a control condition (standard paper road map with only a destination address provided). To obtain voice instructions, drivers operated a tape recorder that allowed them to play either the next instruction or a repeat of the last instruction. Roughly one set of voice instructions per turn were provided. Only information available on the customized map was included in the voice guidance. The customized route maps included only information relevant to the particular route, were drawn to scale, made use of color, included interturn distances, and showed landmarks. The route to be driven was traced in red.

Based on previous research on wayfinding and human memory processes, the authors prepared a template to develop voice instructions that included a sentence to provide some landmarks by which to judge if the driver had gone too far and missed the turn. In addition, the voice instructions always included a summary that reiterated the information about distance to turn, turn direction, and street to turn onto. An example of one set of voice instructions provided for a single turn is provided below:

> *"Drive for one mile to the tricky intersection of Broadway, Norwood Ave., and Bath Ave.*
> *Turn right onto Bath.*
> *The Shadow Lawn Savings and Loan Co. is on the right corner.*
> *If you come to Third Ave., you've gone too far.*
> *Remember, it's one mile to your right turn onto Bath."*

Results were based on 57 test participants driving routes in Monmouth County, New Jersey. Drivers who listened to the voice guidance alone drove less miles and took less time than the control, customized map, or customized map plus voice conditions. Surprisingly, the voice-only guidance involved fewer errors than either the customized map or customized map plus voice conditions. Typical errors were: (1) unable to find location while searching, (2) saw location while driving past, (3) didn't go far enough on road, (4) thought on wrong street, but correct, (5) turned in wrong direction, (6) turned onto wrong street, (7) missed location, but not aware of it, (8) never found correct road. The voice instruction substantially reduced wrong turn direction errors and turns onto the wrong road. Many drivers disliked the "too far" voice instruction, though this may have curbed instances where drivers drove past their turn points.

Streeter et al., (1985) concluded that the auditory channel for directional information appears to be ideal on many counts. They also emphasized that the form and wording of the voice instructions

were the result of "designing, testing, and iterating." They did not claim they were the best possible directions, only that they appear to work effectively, and require information that might reasonably be expected to be contained in a good geographical data base. This is a landmark study of purely auditory route guidance instructions and was the inspiration for numerous other attempts to build auditory route guidance interfaces. Unfortunately, the assessment did not examine other facets of driver workload that also relate to safety such as driver visual allocation, speed variation, lanekeeping, and so forth.

The first human factors evaluation to explicitly address the attentional demands of an electronic route guidance system that provided an electronic map display was conducted on the ETAK Navigator® by researchers at Virginia Polytechnic Institute and State University in the late 1980s (Dingus, Antin, Hulse, and Wierwille, 1989; Antin, Dingus, Hulse, and Wierwille, 1990). The ETAK Navigator driver interface consisted of an electronic map display presented on a 7.3 cm high by 9.5 cm wide monochrome CRT monitor mounted in the center of a test vehicle dashboard. The system used a form of dead reckoning and a map database to provide a dynamically updated heading-up electronic map display of the host vehicle position and the location of the destination. The Etak Navigator did not provide a recommended route; it only showed the destination and current vehicle location. Antin, et al. (1990) point out that such a system requires the driver to perform trip planning tasks in transit instead of pre-drive. The electronic map format, combined with small-screen displays used for automotive applications, also requires the driver to manipulate the electronic display with pan and zoom.

Experimentation was performed to compare driver performance and behavior while navigating with the Etak Navigator moving map to use of a conventional paper map or navigating a memorized route. From the standpoint of visual demand, the moving map display significantly drew the driver's gaze away from the driving task relative to the norm established in the memorized route condition, as well as in comparison to the paper map. No differences as a function of navigation condition were found in brake, accelerator, or lane deviation measures. Further, there were no crashes or near misses during the course of the study. Antin, et al. (1990) reported that there were fewer steering holds (an indication of intrusion on the driving task) with the moving map display relative to the paper map. Dingus, et al. (1989) presented comparative data on the visual demands of Etak Navigator transactions as well as conventional tasks using dashboard instrumentation. In this way, the demand of the navigation tasks could be related to the demand of common tasks. It was determined that most of the Etak Navigator tasks were within the range of visual demand imposed by conventional tasks using dashboard instrumentation. Note that the Etak Navigator made no use of auditory displays.

Walker, Alicandri, Sedney, and Roberts (1991) evaluated 3 audio (female voice) and 3 visual route guidance interfaces, of low, medium, or high complexity, in the Federal Highway Administration's (FHWA) HYSIM simulator, along with a strip map with written directions. The simple auditory display announced "Left, Left, Left." the block before the turn. The medium complexity auditory display announced "Take 3rd left..." three blocks from the turn, "Take 2nd left...", two blocks from the turn, and "Take 1st left..." one block from the turn. The high complexity auditory display was much more verbose and informative, e.g.,: "You are heading Northeast on Jefferson. You are passing Bellvue on the left. The next four streets are Concord, Canton, Helen, and Grand. Your next

turn is a left (Northwest) onto Cadillac in approximately 1.7 miles.". The complex message was presented periodically along the drive.

The simple visual display was composed of left or right arrows set on 90-degree angles with vertical stems and displayed the block before the turn. The medium complexity display was text displayed on a video screen of the form: "Take 3rd right ➜", three blocks from the turn, "Take 2nd right ➜", two blocks from the turn, and "Take 1st right ➜" one block before the turn. The high complexity visual display was an electronic map display on a 4.9 in by 5.5 in screen. In this display, no route was provided; only the vehicle location on the map and icon of the destination were available. Voice displays and visual displays were never combined in this study.

A total of 126 test participants drove in the HYSIM simulator for the study. The difficulty of the driving task was manipulated by varying crosswinds, presence of another vehicle in the adjacent lane, vehicle gage monitoring, reduced lane width, among other manipulations. Data collected included speed, navigation errors, mean lateral placement, lane position variability, heart rate, and reaction to gage changes.

Results indicated that the audio systems were related to fewer navigation errors for all three levels of complexity. The largest difference in navigation errors between visual and auditory systems was in the low complexity case; the smallest difference was in the medium-complexity case. Drivers using the complex navigation devices drove more slowly than those using less complex devices and this effect was especially pronounced in older drivers (55 years or older). Furthermore, several test participants commented that the complex devices (both auditory and visual) gave too much information. This suggests that there is a benefit to streamlining the nature of the navigation system output (both auditory and visual) to include only essential information.

More recently, McKnight and McKnight (1992) examined the attentional demands of in-vehicle navigation systems on drivers in an open-loop driving simulator. Drivers responded to videotaped traffic scenes through simulated vehicle controls while receiving route information through five alternative in-vehicle displays presented on a 14-inch VGA color monitor: an area map, a strip map, a strip map with a cursor to indicate current host vehicle position, arrow displays preceded with an auditory signal to alert a change in arrow direction, and a composite display of the last two alternatives. Based on data collected from 150 test participants, the results indicated no significant differences among the display alternatives in regards to responsiveness to traffic incidents, an outcome attributed to the relatively small amount of time spent looking at the displays. Older drivers required more time to gather information and anticipated turns less well than younger drivers, but were more responsive to traffic safety conditions. The unusually large system display and the open-loop nature of the simulator suggest a need to validate these results in real world driving conditions.

Parkes and Burnett (1993) compared the attentional demands associated with visual and audio route guidance displays by means of an on-the-road study. Sixteen test participants each drove two trial routes, one with only visual information provide by an LCD screen located in the dashboard along the vehicle centerline at steering wheel height, or with both visual and auditory information. Results indicated that drivers spent less time looking at the visual display and more time looking at the road when audio information was presented compared with when audio information was not presented. No data were collected on an audio-only display.

Fairclough, Ashby, and Parkes (1993) presented results of an on-the-road study that compared the effects on visual allocation of LISB/ALI SCOUT and TravelPilot route guidance systems. The LISB/ALI SCOUT system was comprised of an LCD display which showed direction arrows, lane recommendations, vector distance to the destination and a countdown to a guided maneuver. LISB/ALI SCOUT visual displays were also supplemented by an alerting tone and digitized male voice commands. The LISB/ALI SCOUT system provided a recommended route and turn directions. The TravelPilot system used a CRT screen (green symbols on a black display) which displayed a moving map display with host vehicle icon; it did not recommend an explicit route but rather noted the destination on the electronic map and the vehicle location on the road network. Thus, the routes taken could not be equated across the two systems. For the data reported, however, the data were equated for a destination with comparable road types and traffic density.

Twenty-four test participants navigated through Berlin and drove to two destinations, one with each of the two systems, counterbalanced across all the test participants. Based on driver eye glance data, the results indicated that the LISB system was associated with lower percentages of glance duration to the device (9.2% of the glance durations) as compared to the TravelPilot (12.9% of the glance durations). Significantly more time was spent with eyes on the roadway ahead with the LISB system than with the TravelPilot system (80.5% versus 76.1%, respectively). In terms of the percentage of glance frequencies, there were a greater percentage of glances to the TravelPilot display than to the LISB visual display (27.7% versus 23.8%), and a greater percentage of glances to the rearview mirror and left-side view mirror. It is not clear to what extent the differences reported were due to the greater complexity of the TravelPilot display, the audio display that accompanied the LISB system, or the fact that the LISB display provided an explicit route to follow while the TravelPilot did not.

Burnett and Joyner (1993) reported on an experiment conducted in a real road environment with 24 volunteers using an instrumented car. Each test participant drove two different routes, one with an electronic map-based route guidance system and one route with a baseline method. One baseline method was a "backseat driver", i.e., a human who provided voice instructions; the other baseline method was a highlighted paper map. No details about the nature of the voice instructions was provided in the paper.

Results indicated that test participants made no navigational errors (i.e., did not stray off the designated route) with the vocal instructions. They made more errors with the electronic route guidance system than with the marked paper map. The percentage of glance duration to the route guidance system was greater (20.4%) than when glancing to the marked paper map (5.3%). The percentage of total time spent looking to the road scene ahead was greatest for the vocal instructions condition (91%), followed by the marked paper map (85.8%), and lowest for the electronic map display (72.4%). Also, the percentage of total time spent glancing at the rear view mirror (2.1%) and dashboard (0.9%) was less with the electronic map display as compared to the vocal instructions (3.3% and 1.3%, respectively).

Map displays are generally hard for the driver to use while driving (see below). However, if an electronic map is to be displayed, its orientation may make its use while driving relatively harder or easier. The orientation of map displays (e.g., north up, heading up, heading separated) has been investigated experimentally (e.g., Prabhu, Shalin, Drury, and Helander, 1996). Results generally support a heading up orientation for route guidance, north up orientation for route planning.

Large scale operational tests of route guidance systems, funded under the auspices of the Intelligent Vehicle Highway System (IVHS) initiative, began with the TravTek evaluation in Orlando, FL between 1992 and 1993. One aspect of this evaluation was a camera car study by Dingus, McGehee, Hulse, Jahns, Manakkal, Mollenhauer, and Fleischman (1995). In the camera car study, eighteen (18) visitors and twelve (12) local persons of varying ages and both genders served as test participants. Four different navigation system conditions and two control conditions were examined:

- turn-by-turn guidance screens with voice guidance;
- turn-by-turn guidance screens without voice guidance;
- electronic route map (heading up for guidance, north up for planning) with voice guidance;
- electronic route map without voice guidance;
- text directions on a paper list (14 pt Times Roman font);
- conventional paper map.

The TravTek turn-by-turn with voice condition and the paper text directions list provided the best overall performance. The turn-by-turn without voice condition and the electronic map with voice were comparable in many respects but did not perform as well in terms of driving performance and safety-relevant driver errors. The electronic map without voice condition was the least safe of all the conditions test. However, driver performance improved with TravTek navigation system experience. Older drivers (65+)consistently showed lower navigation performance, longer eye glance times, and longer planning and trip times. Older drivers in particular benefited by the TravTek turn-by-turn with voice navigation assistance and had the most difficulty with paper maps. The paper map condition was the least usable navigation approach and resulted in substantially worst navigation performance than any other conditions. Kimura, Marunaka, and Suguira (1997) reported similar findings. Average glance duration and number of glances per minute to a route guidance system display were substantially reduced when voice was present as opposed to when only the visual display was available.

More recently, other operational field tests have been undertaken to examine driver-vehicle performance with route guidance systems. These include the ADVANCE evaluation completed in the Chicago, IL area and the FAS-TRAC evaluation currently underway in Detroit, MI. While the results of such studies have yet to be published, they represent further evaluations in the spirit of TravTek. McGehee (personal communication, 1997) has indicated that the ADVANCE field operational test yielded some interesting results. First, the advantage found for voice displays found in the TravTek evaluation was not replicated in ADVANCE. However, the ADVANCE voice display did not provide street names while the TravTek display did and this may be the chief reason for the lack of benefit. Second, the ADVANCE evaluation also included printed text directions as a comparison condition and again these text directions were among the best of the route guidance presentation formats evaluated. This suggests that printed text directions can be great utility in driver navigation yet commercially available route guidance systems do not make use of text displays. Such displays can be demanding of driver visual attention (Tijerina, Kiger, Rockwell, and Tornow, 1995), but appropriate design may render text displays highly effective and workload reducing.

Srinivasan and Jovanis (1997) have most recently reported on the effects of selected route guidance system interfaces on driver reaction times in a fixed-based, visually high fidelity simulator. Drivers

navigated a simulated network using each of five different route guidance systems: paper map, head-down turn-by-turn display, head-down electronic route map, head-up turn-by-turn display, and an audio-only guidance system. The head-down electronic route map was shown in heading-up orientation on a 15.38 cm (approximately 6 inches) liquid crystal display (LCD) located in the instrument panel to the right of the driver's forward field of view. The route was shown in red and other segments were in green, along with an arrow icon showing the current vehicle position. Prior to the start of the drive, the complete route and the network were shown on the display During the drive, the map scale was changed to half-mile scale. A paper map condition involved a paper version of a full-scale electronic map with route highlighted in pink, other roadway segments shown in green, of size 28.2 x 43.5 cm (approximately 11 x 17 inches). Head-down turn-by-turn displays also made use of the LCD and consisted of the speed in miles per hour, name of the street on which the host vehicle was traveling, name of the next decision street (i.e., street to turn onto), and the distances in tenths of miles or feet to the next decision point. A vehicle icon (amber in color) moved up the currently traveled street (presented as a vertical green line with an arrow head on top). As the vehicle approached the decision point, the distance indicator changed also. Beyond 122 m (400 feet), an amber triangle appeared to indicate whether the driver should turn left or right. At 122 m, a large green arrow with an elongated tail replaced the small amber triangle and units of distance displayed changed from tenths of miles to feet. The display also indicated the geometry of the intersection. The head-up display used the same format but was presented as a virtual image in the simulator focused approximately 2.4 m (7.9 feet) from the driver's eye and positioned in front of the driver just above the hood line on the roadway scene. Finally, voice guidance consisted of a prerecorded female voice that presented two messages per turn. A "distant" message was presented that said, e.g., "In 400 feet, turn right onto Zuma." An "at turn" message, presented at 61 m (200 feet) from the turn, said, e.g., "Turn right onto Zuma." No landmark information was provided, nor was the driver able to request vocal updates on demand.

Reaction time to a scanning task was the response variable analyzed. The test participant was asked to monitor coral-colored squares of an outline form that were presented on the left and right edges of the roadway approximately 28 feet apart. Periodically the squares would turn 45 degrees and become diamonds. As soon as drivers detected the change, they were to push a button on the left or right of the steering wheel hub, depending on which square had turned to become a diamond. The diamond shape remained on the screen for 5 seconds or shorter, depending on the driver's response. It should be mentioned that to increase the driving task demand in the simulator, vehicles from a side street would intrude into the test vehicle's travel lane, pedestrians would stand or walk beside and into the intersections, traffic control signals would change from green to yellow to red, stop signs were present for the driver to obey, and other traffic (described as low traffic density) was present on the simulated roadway.

Results indicated the following. When using the audio system, the drivers response times were fastest while they were slowest with the paper map. The head-up turn-by-turn display was associated with slower reaction times than the identically designed head-down turn-by-turn display. The head-down electronic map was, surprisingly, associated with faster reaction times to the scanning task than the head-down turn-by-turn display. The authors reported that exit interviews revealed drivers liked knowing the number of blocks to a decision point, information that could be derived from the map display but not the turn-by-turn display. Perhaps this accounted for the surprising result. Also, some test participants found the audio display annoying and wanted to have the option of turning it off.

This study indicates that an audio-only display system allows the driver to notice changes in the forward field of view more readily, and that a head-up display (HUD) also increases the speed with which a driver may notice changes in the road scene. However, the outcomes in the simulator with the scanning task must be validated in future studies that examine the driver's ability to detect, perceive, decide, and react to real world objects and events in a real-world driving environment.

These studies taken together suggest the following pattern of results. First, maps are difficult to use for many people and electronic maps are as well. Systems that provide a recommended route can be less workload-intensive. Turn-by-turn displays are less demanding on the driver and support good wayfinding performance. Voice displays can be very effective in route guidance, but their benefits depend on their content. Text directions have proven to be very effective at providing route guidance information with lower driver workload than electronic maps or paper maps. Finally, landmarks are commonly used by people to navigate, yet are conspicuously absent from commercially available route guidance systems even though databases exist that can support the use of landmarks.

1.3.4 Route Guidance Systems Interfaces: Modes and Codes: As indicated in the previous section, a substantial amount of research has been conducted on different ways to present route guidance information to drivers while they are en route. In terms of visual displays, paper maps of various sorts and electronic maps have been investigated for their workload impacts, as have various means to enhance their presentation. An alternative to the map display is the turn-by-turn display which uses alphanumerics (e.g., street names, distance remaining) and special symbology (e.g., turn arrows, distance-to-turn bars, GPS reception icon, heading indicator arrow) to indicate the street being driven, turn direction, distance to turn, street name to turn onto, among other information items. A third type of visual display is written directions, such as those offered by the Hertz Rental Car Company's computerized directions system.

Auditory route guidance displays have been investigated and developed as well. The content of the auditory displays has varied considerably in various research studies. Content ranges from turn-direction only, to distance to turn plus turn direction, to systems or prototypes that also provide street names, and landmarks. In general, the very terse auditory displays have been used in conjunction with visual displays of either the map or turn-by-turn variety. It is noteworthy that Streeter and Vitello (1986) reported that all test participants in their study reported being able to follow verbal directions well, regardless of their spatial or map reading abilities. Also, all test participants reported relying heavily on landmarks for navigation. The researchers concluded that verbal route guidance instructions represent the lowest common denominator for communicating navigation information. As will be seen below, both theory and subsequent research have tended to support these conclusions.

What reasons are there to believe verbal (either spoken or written) route guidance should be beneficial in reducing driver workload? Wickens and Carswell (1997) succinctly characterize several reasons why verbal text (and, by extension, vocal) directions (termed 'route lists' by them) should be very effective as route navigation aids. First, they eliminate the need for many spatial cognitive transformations that may arise with maps. That is, verbal directions, if properly presented, eliminate the need to rotate a map to a heading-up orientation, to pan or zoom in or out of a map display, or to align a 2-D map to the 3-D view of the world obtained while the driver looks through the windshield. Second, the verbal navigational aid provides instructions in terms of a language of

actions (e.g., "Go 0.2 miles and turn right at the Burger King onto High Street"), whereas a map (even a map with highlighted routes and icons to indicate the both driver and destination locations) requires the driver to interpret the spatial information into a set of actions. Third, because the route lists are presented verbally (through voice or text), they can be mentally represented in working memory as a phonetic or verbal code. The multiple resources theory of Wickens (1992) predicts that keeping a driving instruction in working memory as a verbal code reduces competition for the spatial visual processing resources that are involved in many aspects of scanning the driving environment and controlling the vehicle.

1.3.5 Verbal Route Guidance as a Means to Reduce Driver Workload: Most route guidance systems evaluated in the literature (TravTek being a notable exception) provide auditory output that indicates only direction of turn (e.g., "Left turn ahead" rather than "In 0.5 miles, turn left onto High Street" or " In 0.5 miles, turn left at the Pizza Hut onto High Street"). Dingus and Hulse (1993) explain that most voice displays have been implemented as limited-vocabulary digitized sampled speech to keep down costs and enhance message comprehension. Sampled speech has generally sounded better than synthetic speech, but recording thousands of street names has not proven attractive from the standpoint of cost. Recent technological breakthroughs are making voice synthesis systems lower in cost and easier to understand. This opens the way for including more than just direction of turn information in a route navigation system voice message. Furthermore, databases exist that can provide landmark information about commercial establishments, and major geographical landmarks such as railroad crossings, hospitals, parks, cemeteries, and the like. The merging of such database information into route guidance databases is both feasible and has already begun. The DeLorme Street Atlas® for example contains details about geographical landmarks such as railroad crossings, parks, and bodies of water.

As indicated above, three findings seem to have been underutilized to date in route guidance system design. First, voice-only route guidance systems have the potential to work well and eliminate the visual demands associated with a video display. Second, written directions were also reported to be at the top of effective route guidance display concepts. Third, landmarks are considered useful by most travelers. A fruitful line of research, therefore, would be to investigate the properties of voice directions and text directions that maximize wayfinding performance while minimizing driver workload.

Means, Fleischman, Carpenter, Szczublewski, Dingus, and Krage (1993) developed a very sophisticated voice display for the TravTek auditory interface. The system had buttons on the steering wheel that could be used to request auditory information from the system such as Where am I?"(Gives the nearest cross street then the current street being traveled), "Repeat voice" (Repeats last message verbatim preceded by 'The last message was...'), requests for traffic reports and a button to enable/disable the auditory route guidance. Also, the system would no longer repeat an old message after an unspecified "short period" of time. Instead, the message was repeated with "No recent message to repeat". Each maneuver might utilize up to three communications to the user. A "far away" message gave only the distance to the next turn (it did not include the direction of the turn). A "Near turn" message gave distance to turn, direction of turn, and street name. This corresponded with a change in the visual display, which depicted the geometry of the maneuver and street name. Finally, there was an "at turn" message that repeated the near turn message but omitted distance information. Examples of each of the three message types were provided below:

Far away message:	"Ahead, next turn in three and four-tenths miles."
Near-turn message:	"In eight-tenths miles, turn right onto the ramp to I-4 East."
At-turn message:	"Turn right onto the ramp to I-4 East."

Several design goals guided the TravTek auditory display implementation. One was to minimize unnecessary "chattering" to overcome negative reactions to earlier implementations of "talking cars". Unnecessarily verbose messages should be avoided. The authors suggested that anthropomorphism is inevitable when using a voice display, but designers should take steps to minimize its effect. Using a synthesized (computer generated) voice rather than a digitized human voice was recommended.

A second design goal is to maximize intelligibility, in part by allowing the user to adjust relevant settings; the voice synthesis system used in TravTek study allowed the programmer to adjust speed, pitch and voice gain. The TravTek system used a voice resembling that of a medium-sized male because of hardware constraints more than ergonomic reasons. The best voice gender to use is unclear, but some research suggests that a female voice is less intelligible than a male voice. Additionally, since spelling is a poor predictor of pronunciation in English, pronunciation algorithms often fail to derive the correct pronunciations, so the authors suggest checking the system's entire vocabulary. The authors checked, and corrected when necessary, the pronunciation for the entire vocabulary, including 12,000 street names. The authors made two additional suggestions to improve the intelligibility of synthesized speech. First, include the street name suffix (i.e., "Road", "Drive", etc). Second, add an alerting preface (e.g., beep) to attract the driver's attention prior to the voice onset.

A third design goal was to provide timely, useful information in the auditory display. Means et al. (1993) note that in route guidance, each turn instruction should include the street name, direction of turn and distance to the turn. Street names are especially useful at complicated intersections and closely placed streets and are a safeguard against system errors (GPS errors, database errors, etc.). However, one problem with using street names is that the signs are not always visible; here notable landmarks may be quite beneficial. Distance to the next turn should be given in unambiguous units (i.e., miles or kilometers), rather than more ambiguous units such as "go two blocks" which may be ambiguous when the cross street intersect the current road on one side. They argue that stating the time to turn as an objective distance (e.g., "Turn in 1.4 miles") is useful for people who can gauge the distance, and not harmful to those who cannot. The authors do not mention dynamic updates, which probably makes this problem irrelevant. For example, the TravTek video screen could show the dynamically decreasing distance to the turn. Similarly, the "Where am I?" button gave the upcoming street followed by the current street in a timely fashion. By pressing the button repeatedly, the system would output at each cross street that the driver approaches. It is interesting to note that despite the advanced capabilities of the TravTek voice display, it was never tested alone (it was always paired with a video display of some type) and landmarks were not incorporated into the design.

Kintsch and Van Dijk (1978) have presented a model of text comprehension that suggests the complexity of a sentence is a function of the number of underlying "propositions" or ideas it contains. They estimate that only four ideas or propositions can be held in working memory at one time. In the case of verbal route navigation directions, this suggests that if only four propositions

can be reliably held by the driver in working memory, then a driving instruction should contain no more than the following: a) the distance to the next turn, b) the turn direction, c) a landmark that designates the turn, and c) the street to turn onto. Note that these information units or propositions were also contained in the recommendations of Streeter et al. (1985), and Green, et al., (1995) for auditory voice displays.

Regarding the presentation of text displays, some basic research suggests a simple way to enhance text comprehension. Graf and Torrey (1966) reported that comprehension of text was enhanced when sentences were broken into several different lines of text and the end of each line corresponded to the end of a phrase. They presented sentences to test participants a line at a time. The passages could be presented in the Form A, in which each line is a major constituent phrase, or in the Form B, in which this is not so.

Form A	Form B
During World War II, even fantastic schemes received consideration if they gave promise of shortening the conflict.	During World War II, even fantastic schemes received consideration if they gave promise of shortening the conflict.

Test participants showed much better comprehension of passages in the Form A, presumably because the physical separation of lines coincided with the end of phrases. An extrapolation of this research suggests that instructions that must appear on several lines (or as few words on successive screens) should be divided by phrases rather than, say, based on line length. Thus, it is predicted that "Watch your step....when exiting.... the bus" will be understood more quickly than "Watch your...step when....exiting the bus" (Wickens and Carswell, 1997). This finding can be applied to the preparation of text directions for route navigation as well.

These concepts could be developed and prototyped for predetermined origin-destination pairs that drivers would traverse during data collection in an instrumented vehicle. The relevant workload measures of merit would include driver eye glance behavior measures, driver-vehicle performance measures, global wayfinding measures, and subjective assessments. The results of this work will augment current DOT human factors guidelines for route guidance systems or Advanced Traveler Information Systems (ATIS), guidelines which currently do not sufficiently address either voice displays or text displays for wayfinding. Safety relevant outcomes might be more efficient visual scanning with less eyes-off-road time, less en-route speed variation, fewer abrupt or aggressive lane changes to position the vehicle for a sudden turn, fewer inadvertent lane exceedences, lower travel time (affecting exposure), and the like.

As indicated in the literature review, there is evidence from prior research that people can work well with textual/verbal directions when wayfinding, regardless of their spatial abilities. There are reasons, both theoretical and empirical, to believe that verbal guidance will interfere less with the primary driving task than most forms of spatial guidance. While it is expected that a text display will be associated with some degree of visual demand (say, relative to a voice-only display), perhaps

it will be sampled infrequently and provide drivers a sense of perspective (where they have been, where they are going) and useful information in the form of distance-to-turn information. Zaidel and Noy (1997) reported that voice guidance, with context-rich information (landmarks and other orientation cues) was more effective than displaying guidance on a visual display. However, they also found that an automatic voice-only display system produced higher subjective assessments of driver workload, perhaps because of the demands such displays placed on the driver's short-term memory. It is interesting to note that in their Experiment 1, the "ideal navigator" display consisted of a paper display of redundant text directions together with automatically presented voice guidance that included landmark, geometry, and distance information.

1.3.6 The Use of Landmark Information in Route Guidance Systems: Why should landmarks in route guidance instructions be beneficial to the driver and reduce workload? Streeter and Vitello (1986) noted that all test participants in their study reported relying heavily on landmarks for navigation regardless of their spatial or map reading abilities. Except for street names, landmarks are not included in commercially available route guidance systems, despite the fact that databases exist that could provide landmark information. Street signs may be considered the most rudimentary landmark but street signs are generally numerous, similar in appearance, and hard to read at a distance. More visible or notable landmarks may be categorized in terms of man-made features (e.g., traffic lights, traffic signs, businesses, cemeteries, private buildings, fences, railroad tracks), terrain features (e.g., hills, embankments, woods, rock formations, cul-de-sacs), water features (e.g., ponds, culverts, rivers), and vegetation features (e.g., corn fields, gardens, open fields, distinctive trees or landscaping) (Green, et al., 1995; Whitaker and Cuqlock-Knopp, 1995). The differential benefits of different types of landmarks are not fully understood.

Akamatsu, Yoshioka, Imacho, Daimon, and Kawahsima (1997) reported on a study with eight drivers in Tokyo, four of them familiar and four of them unfamiliar with the area. The test participants "thought aloud" as they attempted to find destinations with one of a variety of route navigation system prototypes. The verbal protocols indicated the types of landmarks and incidence of their mention by the drivers. More than half of 246 spoken words referred to structures, street names, and intersection names. This is perhaps not surprising given the route navigation systems indicated the recommended route with such information. The authors noted that frequent use of the navigation system at intersections suggested that drivers used landmarks to decide whether or not to turn.

A very subtle effect of the order in which route guidance information elements are presented in an instruction has been reported by Jackson (1996a, 1996b, 1996c). The three articles report two studies that use a videotaped route to simulate way finding in an unfamiliar environment. Subjects viewed two videotapes for each of three interconnecting routes. Each videotape was made with a video camera fitted with a wide angle lens, allowing the experimenter to control what the subject was able to see along the route. The first video was filmed with the camera pointing in the direction of travel. In the second video, the camera pans from left to right to highlight landmarks. Each of the routes involved a minimum of four turns and shared a portion of the route with the other two routes. The staging area (based on the map) seemed to occupy about 900 sq meters. One study involved the following groups:

Group 1: Subjects watched video in silence.

Group 2: Subjects heard "complete route instructions" (it is unclear what this means).
Group 3: Subjects heard landmark information followed by turn directions
Group 4: Subjects heard turn directions followed by landmark information

After the subject viewed both videos for each route (the "straight-ahead" and the landmark videos), the subject was asked to perform a series of tasks, which are "highly correlated" (no specifics given) to the measure that is reported. The tasks included arranging ten photographs in the order that they appeared along the route, indicating the distance between those scenes, indicating the direction of one photograph relative to another, and estimating the time and distance traveled on each route. After completing all three routes subjects were asked to complete a task similar to one of the above tasks, but in this case subjects indicated the direction to landmarks between routes (cross pointing task). Subjects also drew a sketch of the area. The author reports that the landmarks-then-turn direction instructions elicited the best performance in the pointing task, although this is only the case on the third route and the cross-pointing task, suggesting that subject performance improved with directions given in this order, (but order and route number seem to be confounded). The opposite occurred with the directions/landmark group (i.e., they got worse on the third route). The difference in errors on the third route is about 27 degrees, and about 12 degrees on the cross route pointing task. Jackson explains the difference by noting that while one piece of information is being processed, the other must be stored in working memory. Since the direction is unambiguous, it should be processed more quickly, than the landmark, and thus cause fewer problems with storage/retrieval. Jackson also describes an additional analysis showed that subjects driving less than one year were less accurate than those who had been driving for 2-3 years. This result bears upon the issue of driver cognitive maps, something of concern only if the driver wishes to learn his or her way around in a new locale. However, the wording difference may also be usable to drivers who wish only to get to a previously unknown destination once (e.g., travelers).

Green et al., (1995) recommended a template for auditory route navigation messages that involved early, prepare, and approaching messages (similar to those of TravTek). The position of the turn direction is opposite that recommended by Jackson's research. Incorporating this change, the following message set might be provided:

- Early message: "In 3.5 miles, bear left at the Pizza Hut onto Green Street."
- Prepare (Near-turn) message: "In 1 mile, bear left at the Pizza Hut onto Green Street."
- Approaching (At-turn) message: "Bear left at the Pizza Hut onto Green Street."

The timing of auditory route guidance information presentation has been studied by at least two independent laboratories. Ross, Vaughan, and Nicolle (1997) conducted a study with fifteen test participants in which a passenger (experimenter) presented short, precise navigation instructions that contained an intrinsic prompt (e.g., "Take the next left turn", or " Take the second right turn"). These were presented at varying times and the driver indicated on a 6-point scale whether the instruction was too early/too late. No visual display was used and no landmark or street name information was provided in the voice directions. Based on this approach, the authors presented regression equations that indicated the preferred minimum, ideal, and preferred maximum presentation distance (from the intersection) as a function of travel speed. Green and George (1995) conducted a study of 48 drivers to address the same issue of timing the onset of auditory displays for final turn instructions. In one part of their study, they had test participants approach a known

intersection and say "Now" at the latest moment when the test participant would feel comfortable hearing the turn direction information. Their results are presented in terms of a single regression equation that predicted distance from turn for final turn instructions. This equation included, in addition to travel speed, the predictor variables of driver age, gender, turn direction, and number of vehicles ahead. Their results suggest that last turn messages should be provided approximately 450 ft prior to the turn, with that value being adjusted 15 ft for each mile per hour change. Adjustments are also made for gender, age, and turn direction. Like the study of Ross et al. (1997) the voice guidance did not include either street name or landmark information. It is unclear to what extent the recommended timing would vary with the inclusion of such information in the auditory message. Kimura, Marunaka, and Sugiura (1997) note the 300 meters was generally the distance prior to an intersection where landmarks were identifiable studies conducted in Tokyo and Nagoya, Japan. They also recommended that turn directions be provided approximately 700 m prior to the intersection to accommodate lane change maneuvers that might be required into the turn lane.

1.4 Voice Recognition as a Means to Reduce Driver Workload

The automotive industry is actively working to adopt voice recognition technology. Furthermore, text-to-speech processing is also becoming available for automotive use. The use of this technology is appealing because voice recognition and text-to-speech technology allow the driver the possibility of performing tasks without visual or manual demand. However, even though there is no visual or manual demand, voice transactions may still be distracting to the driver.

Voice technology could allow the driver to control many functions with voice commands that are currently performed manually. Quite a few tasks are possible to perform using voice commands. One application is in voice communications. For example, voice technology would allow hands-free destination entry, cellular telephone dialing, answering and hang-up, hands-free telephone conversation, and voice memos. Other applications would be controlling and monitoring vehicle subsystems such as radio tuning, volume control, CD control, or HVAC controls.

Voice technology could also give drivers more convenient access to information sources such as address books, email and fax text-to-speech readout, or electronic yellow pages. Voice technology could also be used for navigation system functions such as destination entry and wayfinding support. On a lighter note, voice technology also allows the possibility for voice interactive games such as blackjack or trivial pursuit.

Before this technology is implemented, several important questions must be asked. What benefits are accrued by voice transactions and when do they arise? What problems does a given voice transaction create for drivers and under what circumstances? What functions should be available or not available using voice technology? And finally, under what conditions does this technology interact with driver age?

Multiple-Resource Theory (Wickens, 1992) provides some insight into dual task performance. Multiple-Resource Theory suggests that resources available to perform a task may be defined by three dimensions: encoding modality such as auditory or visual, processing code such as verbal or spatial, and response modality such as manual or vocal. Each combination of these dimensions may

be thought of as having a particular capacity or resource availability. Furthermore, if two tasks are characterized by non-overlapping combinations of these dimensions the person engaged in the two tasks could be expected to perform as well performing them concurrently as they would if they performed them separately. For example, most people can successfully walk down a hallway (a visual spatial task with a manual response) while at the same time talking to a friend or coworker (an auditory verbal task with a vocal response). On this basis it might be predicted that it should be easy for a person to drive a car and listen to the AutoPC read back their email or talk to someone on a hands-free cellular phone. As Wickens (in a published letter to the editor) points out, however, this neglects the role of the central processing code. According to the Multiple-Resource Theory cognitive processing can be characterized as either verbal or spatial. It is possible for a visually presented task responded to manually and an auditorally presented task responded to vocally to both draw on the same central processing resource. For example, an architect listening to some changes that have been made to some plans over the phone may draw upon his or her spatial processing resources in order to understand what has been done. If the architect is driving at the same time, this has the possibility of reducing his/her spatial processing resources available to survey the scene out the windshield and make appropriate control inputs. A worse case could be if the architect begins making important decisions and planning changes that could use a combination of verbal and spatial processing, which would be very resource demanding. This situation could draw a significant amount of the architect's attention away from the driving task even if the cellular phone is hands-free and despite the fact that speaking over the phone is an auditory/vocal task. Furthermore, responding over the phone, while it is a vocal response, will be a fairly complicated vocal response.

Wickens (1992) describes the vocal response as a separate response modality, however, it is actually a complex motor response that has the potential to interfere with other manual responses (such as braking or steering). For example, Van Hoof and Van Strien (1997) found that reading a list of words out loud produced a small degree of interference in subjects' finger tapping speed when the two tasks were performed concurrently. They also found that reading speed decreased when the two tasks were performed concurrently. It seems likely that a larger degree of interference may have been found if the subjects had been required to generate and speak aloud more complicated sentences.

Interestingly even walking and talking to someone at the same time may be more difficult than performing each of those tasks separately. Bardy and Laurent (1991) conducted an experiment in which subjects walked towards and stopped in front of a target while listening for a change in pitch of a series of auditory tones. Bardy and Laurent found that although there was no measured impact on walking performance, reaction times were delayed for the auditory task. This result is somewhat surprising given the simplicity of the two tasks. In this experiment, the subjects were to press a button to indicate a change in pitch of the tone. It is possible that allowing a vocal response to the auditory task in this case may have produced less interference.

The literature on dual task performance suggests that performance depends a great deal on the type of tasks being performed and the input/output modalities. McLeod (1977) found that a two-choice tone identification task interfered little with a manual tracking task when subjects could respond vocally but interfered significantly more when the subjects responded manually. Wickens et al (1983) asked subjects to perform a Sternberg memory task, in which subjects had to identify whether or not a target letter was presented in a previously presented group of letters, concurrently

with a tracking task, in which subjects had to pursue a target moving around on a screen with a joystick controlled cursor. In addition, the memory task was presented auditorally on some trials and visually on other trials. Subjects also responded to the memory task either manually or vocally in different trials. The results indicated that the memory task was disrupted by common inputs (performance was worse when both the memory task and tracking task were presented visually) but depended less on the response mode (vocal or manual). The tracking task performance, however, was disrupted mostly by the output modality and depended very little on the presentation mode. Performance was much worse on the tracking task when both the tracking task and memory task were responded to manually, whether the memory task was presented auditorally or visually. Performance was not affected as much when the memory task was responded to vocally but, again, did not matter if the memory task was presented visually or auditorally. It is important to note that for this experiment the letters for the memory task were presented very near the tracking display such that the subjects could view both the tracking display and the letters at the same time without moving their eyes.

In a second experiment Wickens et al (1983) found that a spatial secondary task interfered with the primary flight simulator task more than a verbal secondary task. Furthermore, when the secondary verbal task was presented auditorally and responded to vocally it interfered very little with the primary flight task. When the verbal secondary task was presented visually and responded to manually, however, performance on the primary flight task decreased substantially.

In a more recent study by Sarno and Wickens (1995) the results of a dual task evaluation were similar to the Wickens et al (1983) results in that the primary tracking task performance was affected mostly by the output mode. Performance on the tracking task was worse when the secondary task, whether it was verbal or spatial or easy or difficult, was responded to manually. Performance on the primary tracking task was not as badly affected when the secondary task was responded to vocally. However, an unusual result of this study was that the verbal task interfered more with the tracking task in most cases than the spatial task. According to Multiple Resource Theory the spatial task would be expected to interfere more with the tracking task than the verbal task. Wickens et al (1983) point out that the spatial-verbal distinction is really more of a continuum than a dichotomy. It may be difficult to neatly classify some tasks as either spatial or verbal. In the Sarno and Wickens (1995) experiment the verbal task was subtraction in the easy condition and multiplication in the difficult condition. It is possible that mathematical reasoning is not entirely verbal but is actually a combination of resources.

The dual task studies described above generally indicate an advantage for an auditory presentation and a vocal response to a secondary task relative to visual/manual presentation and response for the secondary task. This is encouraging if voice technologies are going to be used for automotive applications. However, it should also be noted that in every case the performance on the primary visual/manual task was worse when combined with any secondary task, even when the secondary task was an audio/vocal task.

It is becoming clear that cellular phone use while driving has the potential to distract drivers enough to increase the risk of an crash. Redelmeir and Tibshirani (1997) found that drivers' risk of an crash was four times higher when using a cellular telephone compared to the same drivers' risk when they were not using their cellular phones. Significantly, the increase in risk was the same for both hand-

held cellular phones and hands-free cellular phones. Apparently it is not simply the fact that the driver may only be driving with one hand that can lead to an crash. Rather, the important factor may be a loss of attention to the driving task. More direct evidence for interference of driving when using a phone comes from Brown et al. (1969). In this experiment subjects were required to decide whether or not to drive through a gap while at the same time performing a verbal task. The gap size varied from 9 inches wider than the car to 3 inches narrower than the car. The verbal message was transmitted to the subject via the equivalent of a hands-free telephone. The subject did not have any buttons to press and did not have to hold the receiver. The verbal task was a grammatical reasoning task. Subjects had to decide whether or not a sentence describing two letters matched the order of two letters presented following the sentence (for example: B is not followed by A - BA: False). The subjects responded by vocally answering true or false. Brown et al. found that subjects made significantly more mistakes in gap size estimation when concurrently performing the verbal task - subjects were much more likely to try to drive through a gap that was too small. Note that the secondary verbal task in this experiment was an auditory/vocal task yet it still caused interference in the visual/manual driving task. This may be because the verbal reasoning task requires spatial processing resources (picturing the position of two letters), which pulls away needed spatial resources to make the gap size decision.

It is important to note that a person talking on a cellular phone may be qualitatively different from chatting with another person in the car. The driver chatting with another person in the car has the advantage of another pair of eyes watching the road and the further advantage that the passenger will probably pause the conversation should a problem arise that requires the driver's attention. It may also be informative to subjectively rate the ease of listening and conversation complexity for the typical in-the-car versus on-the-phone conversation. The use of a voice technology product such as the AutoPC is very similar to using a hands-free cellular phone and it is possible that it may produce the same degree of interference that cellular phone use seems to produce while driving.

Several factors may affect the impact of voice technology usage. For example, task demand will affect how much interference, if any, the use of voice technology has on driving given that the driver chooses to engage in the voice task. Usage pattern, such as how often the driver uses the voice technology will also affect the degree of interference. The conditions under which drivers choose to use the voice technology will also be important. Another factor which could be significant is that of driver individual difference. It may be that some people are better able to divide their attention between two concurrent tasks than others. For example, Fournier and Stager (1976) found consistent differences in dual task performance among Canadian detection/communications officers assigned to antisubmarine aircraft. Furthermore, this difference was significantly correlated with peer and supervisor ratings of job performance. The officers that tended to perform better in the dual task experiment also tended to be rated more favorably by their peers and supervisors. Kahneman, Ben-Ishai, and Lotan (1973) found consistent differences among bus drivers on a test of selective attention. These differences were significantly correlated with the number of crashes of bus drivers.

As would be expected, amount of driving experience has an effect on dual task performance. Wikman et al. (1998) found that inexperienced drivers were more likely than experienced drivers to deviate in their lanes while putting a cassette in the stereo, changing the radio station, or especially when dialing a phone. Age may also be an important factor in the ability to perform two tasks concurrently. Brouwer et al. (1991) found that older drivers did not perform as well on a lane

tracking task when they had to count visually displayed dots at the same time and respond manually compared to younger drivers. Significantly, however, older drivers performed equally well when the dot counting task could be responded to vocally. This result suggests that voice technology could be helpful for older drivers when they have to perform another task while driving.

In summary, the research on dual task performance suggests that voice technology may provide a means of performing a second task while driving which may produce relatively less interference with the primary driving task than performing the equivalent task without the use of voice technology. An additional benefit of voice technology is that it may be helpful for older drivers when performing a second task while driving. However, all of the research indicates that performing a second task always has an adverse effect on primary task performance compared to performing the primary task alone. The important question is how much interference is there and is it acceptably small? Recent investigations of cellular phone usage while driving provide some evidence that the use of voice technology has the potential to distract drivers enough to increase the risk of an crash (Goodman, Tijerina, Bents, and Wierwille, 1999).

1.5 The Status of Commercially Available Route Guidance Systems

A relative comparison of commercially available systems was recently reported in the trade literature. Paul (1996) reported a product evaluation of the following navigation systems:

- Acura (developed in conjunction with Alpine Electronics)
- Oldsmobile Guidestar (developed by Siemens/Zexel, also the Hertz Neverlost)
- Clarion/Eclipse Interactive Voice Navigation System (built by Amerigon)
- Delco Telepath 100
- Sony NVX-F160.

Characteristics of these systems are represented in Table 1-1. Test drivers in the Los Angeles areas were asked to find three destinations of various types (unknown area, nearest fast-food restaurant, specific tourist attraction). The systems were then rated in terms of ease of route programming, helpfulness in finding a destination, quality of route finding, system versatility, overall ease of use and value. The Acura/Alpine and Oldsmobile Guidestar/Zexel systems were rated the best in terms of convenience and ease of use; the Eclipse Interactive Voice system was rated best in overall value and highest in quality of route finding. No driver performance data were reported. The Sony NVX-F160 was rated poorest in value and ease of route programming, and the Delco Telepath 100 received the lowest rating for helpfulness in finding a destination. None of these systems provided incident detection. While each of the systems were evaluated subjectively, no driver performance-based evaluation was reported. Such an evaluation would be a useful supplement to Paul's (1996) work because subjective evaluations of systems may sometimes be in conflict with the safety or usability of those systems. For example, novelty effects may prompt people to favorably assess a system that in fact is difficult to use. Furthermore, if a system is seen is as an improvement over what has previously been used, there may be a tendency to favorably evaluate a new system that is itself still problematic .

Table 1-1. Navigation System Features for Systems Reviewed in Paul (1996).

	Accura/ Alpine	**Clarion/Eclipse Interactive Voice Navigation (built by Amerigon)**	**Delco Telepath 100**	**Oldsmobile Guidestar/ Hertz Neverlost/ Zexel Navmate**	**Sony NVX-F160**
Navigation Method GPS	✔	Navigation based on digital mapping and origin/destination inputs	✔	✔	✔
Dead Reckoning	✔	" "	✔	✔	✔
Navigation Features On-screen map	✔	No visual display	Provides a compass heading	✔	✔
Turn-by-Turn route programming	✔	✔ Driver must query system periodically (e.g. say "next" after each leg) to receive further instructions	Only compass heading	✔	Does not plot a route; only pinpoints destination on map display and an arrow in direction of destination.
Voice directions	✔	✔ " "	no voice because no directions	✔	no voice because no directions
Voice recognition		✔ Unique system			
Auto. reprogramming for route variation	✔				
Available routing preferences	✔			✔	
Input Method	Hand-held remote control	Voice Input with machine feedback and auditory review of alternatives	Rotary knob/menu scrolling	Four-way scroll button/ menu scrolling	Joystick button on remote control
Database Contents Fine Dining	✔	✔	✔	✔	✔
Family Restaurants/ Fast Food	✔	✔	✔	✔	
Gas Stations	✔	✔	✔	✔	
Hotels	✔	✔	✔	✔	✔
Hospitals	✔		✔	✔	
Banks/ATMs	✔			✔	
Popular Attractions	✔	✔	✔	✔	✔

Recent in-vehicle information technology has attempted to harness the power of enhanced voice recognition technology. The AutoPC represents an example of the most recent such system. Other products include a similar voice recognition and generation system under development at Visteon, as well as technologies being marked by Lernout and Hauspie and others, to name but a few.

1.6 Research Objectives

Clearly, a great many issues arise in consideration of the human factors of wireless telecommunications and route guidance systems for drivers. Based on inputs from staff at the NHTSA, the following objectives were selected for the studies described in this report:

- Characterize the impact of route guidance system destination entry on vehicle control and driver eye glance behavior on a test track;
- Assess the influence of individual differences, as indexed by a battery of cognitive tests, on the susceptibility to distraction as indicated by disruption in vehicle control and driver eye glance behavior during destination entry and cellular telephone use while driving;
- Examine the validity of a proposed SAE recommended practice to assess whether or not a given route guidance destination entry function ought to be allowed while the vehicle is in motion;

1.7 Organization of the Report

This report consists of four sections. The first section is this general introduction to the topic area. The second section describes what constitutes the first route guidance system destination entry study conducted on a test track as opposed to a simulator. Cellular telephone dialing and destination entry with a commercially available voice recognition system are included in the evaluation. The third section presents a study of individual differences and their correlation with driver lanekeeping and eye glance behavior while engaged in various in-vehicle tasks on a test track. The report concludes with a section that describes a study of the so-called 15-second rule which lies at the heart of the proposed recommended practice SAE J2364.

1.8 Summary

Many important human factors consequences of wireless telecommunications and route guidance system use while driving have been introduced here. Many remain to be investigated. Recent trends in technology suggest that the future of in-vehicle "telematics" will involve products which fully integrate telecommunications, route guidance, and other driver information systems functions, perhaps with a voice recognition-based driver interface. The studies and analyses described in this report represent the authors' attempts to contribute to a better understanding of the complex relationship between the driver and the technology. It is hoped that this understanding, in turns, supports a more driver-centered evolution of the technology.

2.0 DRIVER WORKLOAD ASSESSMENT OF ROUTE GUIDANCE SYSTEM DESTINATION ENTRY WHILE DRIVING: A TEST TRACK STUDY

2.1 Introduction

No published data exists on the demands of route guidance system destination entry while driving. Many believe that destination entry while driving is simply too distracting to be carried out safely, but many commercially available systems allow it. Some commercially available route guidance systems provide cautions to avoid distraction while driving but do not lock out such functions while the vehicle is moving. Others systems include a motion sensor that locks out such functions, but this is the exception rather than the rule. Still other systems provide no cautions or lockouts at all. Green (1997) has pointed out several scenarios wherein destination entry or retrieval en route might be attempted: the driver is in a hurry, knows the general direction in which to start, and adds the destination information later; the route guidance system does not use congestion information and a radio announcement indicates the current route is problematic; the driver enters the wrong destination initially and does not wish to stop the vehicle to correct it; or the driver does not know the exact destination at the beginning of the trip, enters an interim destination known to be close by, then enters the actual destination at a later time.

The objective of this study was to examine four commercially available route guidance systems, representing alternative destination entry and retrieval methods, in terms of driver visual allocation, driver-vehicle performance, and driver subjective assessments. No study has examined this problem in a real world driving context, in part because of the very safety concern that prompts an interest in the topic. Therefore, a test track study was conducted with light traffic present and a vigilant ride-along observer in the test vehicle.

2.2 Approach

2.2.1 Test Participants: Sixteen (16) test participants were recruited from the Transportation Research Center Inc. pool of entry-level test drivers in equal numbers of males and females in each of two age categories: Younger (35 years or younger) and Older (55 years or older). These drivers were hourly employees with valid driver's licences and generally less than 2 years of TRC driving experience. None of the test participants owned or had significant prior experience with route guidance systems prior to this study.

2.2.2 Test Vehicle and Instrumentation: The test vehicle was a 1993 Toyota Camry, equipped with MicroDAS instrumentation (Barickman, 1998) which captured travel speed, lane position, and lane exceedences, as well as video of the road scene and driver eye glance behavior at a sampling rate of 30 Hz. Eye glance video was later manually reduced.

2.2.3 Route Guidance Systems: Four (4) unmodified, commercially available route guidance systems, each with a different destination entry and retrieval logic and driver interface, were used in the test. The dash mounted Delco Telepath 100® consisted of a 3-line LCD display to present menu items, scrolled by means of a bezel-mounted rotary knob and selected by pressing an Enter

key. The Alpine NVA-N751A® incorporated a free-mounted 5.6 inch active matrix color display without bezel keys. It displayed an alphanumeric keyboard and entries were made by scrolling from key to key with a joystick mounted on a remote control unit; pressing down on the joystick registered a character or selection. If sufficient alphanumerics were entered for the system to estimate candidate destinations, these were presented as an alphabetized scrolling list of 3 items at the bottom of the display of the alphanumeric keyboard screen. The Zexel Navmate® consisted of a free-mounted 4 inch diagonal full color LCD screen with a set of bezel control keys, including a central "left, right, up, down" key and an Enter key. Both the Zexel and Alpine systems were mounted on a gooseneck pedestal bolted to the floor board between the driver and passenger. The Zexel system presents menu options for destination entry type and city, followed by a scrolling display of numerically and alphabetically arranged destinations generally presented 11 to13 lines at a time. The driver presses the Enter key to make a selection. Finally, the dash mounted Clarion Eclipse® Voice Activated Audio Navigation (VAAN) system used voice recognition and output exclusively; there was no visual display. Keywords would activate the VAAN for destination entry. Destinations were entered by spelling them. The VAAN emphasized precise spelling of a destination; each letter uttered by the driver would be proceeded by a beep to acknowledge receipt of the input. The driver uttered "verify" to conclude an entry. The system would eventuate in a spoken list of best-guess candidate destinations for selection by the driver via YES or NO verbal responses.

The last three of these systems allowed for entry of a street address, intersection, or point of interest (attraction, restaurant, hotel, etc.). Thus, three types of tasks (address, intersection, point of interest) were included as suitable for comparisons among the systems. The Delco system only supported point of interest selection. Also, two additional tasks were included for comparison purposes: tuning a radio to a specific band and frequency using a modern "Seek" function on the Clarion Eclipse system; and manually dialing a cellular telephone (a 10-digit number on a handwritten note card) using a cordless AUDIOVOX Model MVX-500.

2.2.4 Test Route: The TRC 7.5-mile multi-lane test track is in the form of an oval with banked curves at either end and with unbanked straightaways that measure approximately 2.0 miles each. The test track is comprised of three 12-ft wide concrete lanes with a fourth inner blacktop lane for use in the event of vehicle breakdowns or required stops. The test vehicle for this study operated in lane 1 (adjacent to the innermost blacktop lane) and changed lanes only as needed for normal track operations and safety. The test participant was asked to drive at approximately 45 mph on the straightaways and accelerate to 60 mph on the curves, provided that any requested tasks are completed by the time the test vehicle enters a curve. Otherwise, the driver was to maintain 45 mph and attempt to complete the requested in-vehicle task. Traffic density tended to be light relative to open road driving. However, travel speeds for other vehicles of the track might vary greatly, vehicles involved with other testing could slow, stop, or move to the blacktop lane abruptly, and track repair and roadside obstructions had to be avoided. Faster traffic drove on the outer lanes of the oval. Data collection was scheduled for between 8:00 am and 4:30pm weekdays.

2.2.5 Independent Factors, Dependent Measures, and Study Design: A two-between, three-within mixed factors experimental design was used for this study. The between-factors were Age category and Gender. The within factors were: Route Guidance System (Zexel, Clarion Eclipse VAAN, Alpine NVA-N751A, and Delco Telepath 100); Destination Category (Street address, Cross street, or Point of Interest), and Destination Targets (Target 1, Target 2, different for each

destination type but the same targets across each route guidance system). In addition, two non-destination entry tasks were included for comparison: dialing an unfamiliar 10-digit number on a cellular phone and manually tuning a radio to a specific frequency on the AM and FM bands. The dependent measures of interest for this study were: Visual Allocation (mean glance duration, mean glance frequency, and total glance time to road ahead, in-vehicle device, and note card); Driver-Vehicle Performance (number of lane exceedences, lane exceedence duration); and Trial Time (i.e., destination entry task completion time). Driver preferences and impressions of safety were also collected, among other subjective assessments.

2.2.6 Procedure: Prior to the data collection runs, the test participant signed an informed consent form (see Appendix A) and the experimenter familiarized the test participant with each navigation system. Each test participant then completed 12 practice data entry tasks per system (four for each destination category), entered while the vehicle was parked. This training was done in two phases (morning and afternoon); so, two systems were reviewed prior to each half of the test track trials. On the 7.5 mile track, the order of trials were counterbalanced across the four route guidance systems (Zexel, Alpine, Delco, and VAAN), destination entry category (point of interest, intersection, and street name targets), and target (Target A or Target B within a category). All trials with a given system were executed before moving on to another system; the destination type and targets within destination type were counterbalanced to control for order effects. The cellular phone and radio tuning tasks were interspersed between destination entry trials on an opportunistic basis by the experimenter in a quasi-random fashion. Prior to leaving for the test track, the destinations were presented to the test participant in 18-point Times Roman font and the test participant was asked to write in his or her own hand each destination on a separate index card, as well as the 10-digit unfamiliar telephone number, such that they would be able to read from it while driving. A task began when the ride-along experimenter gave the driver a hand-written card or a radio tuning task was requested orally by the ride-along experimenter. The task ended when the request had been fulfilled, as indicated by an event marker triggered by the experimenter. Requests for tasks were generally made when the test participant was exiting a curve onto a straightaway segment of the test track. After test track data collection was completed, the test participant answered the subjective assessment questions and was released.

2.2.7 Data Analysis: The data were analyzed by means of the analysis of variance for split-plot designs using the SAS® Proc GLM routine, Type III Sums-of Squares. Prior to ANOVAs, appropriate transformations were applied (e.g., log transforms of glance durations, square root transforms of lane exceedence counts) to both normalize the data and stabilize what were often heterogenous variances. Outliers were not deleted from the data set unless they were clearly erroneous (e.g., a verified manual data reduction error for eye glance data).

2.3 Results

Only Point-of-Interest (POI) destination entry results will be presented here. This choice is made because a) all four systems were capable of this type of transaction, and b) the results generally follow the same trends as those for a companion analysis that included destination category (street address, intersection, point-of-interest) but did not include the Delco system due to its limited capability. Since the specific destinations were not meant to be comprehensive, but merely a

methodological convenience, specific target effects are not presented here. All results to be presented and discussed were significant at an $\alpha < 0.05$. All other results are considered insignificant.

Figure 2.1 shows the effects of the different systems and tasks in terms of trial time or task completion time. Panel 1A indicates a significant effect of Age on destination entry trial time, with older drivers averaging almost twice as long as younger drivers. Panel 1B shows the average trial times for POI entry as a function of system, with the 10-digit cell phone dialing and radio tuning tasks included for comparison purposes. The longest completion time, on average, was with the Alpine system (118 seconds, approximately), the shortest average completion time was with the VAAN and Delco (approximately 75 and 78 seconds, respectively). Note also that all of the POI destination entry tasks took significantly longer than manually dialing an unfamiliar 10-digit number (approximately 28 seconds) or manually tuning a modern radio (approximately 22 seconds). Panel 1C is significant in that the Age difference is "neutralized" by the use of the VAAN voice data entry system.

Figure 2.2 presents the average number of glances and mean single glance duration data associated with device and note card for this study. A main effect for age found that older test participants made significantly greater numbers of glances per POI destination entry than younger participants (approximately 31 vs. 16 glances, respectively). Panel 2A shows, not surprisingly, that the average number of glances per transaction were trivial for the VAAN in comparison with other route guidance systems, and even lower than the cellular telephone dialing and radio tuning tasks. No interaction between Age and Device was found. Panel 2B, on the other hand, reveals that the VAAN was associated with over twice as many glances to the note card, on average, than any other system. Presumably, the greater precision required to spell the destination correctly prompted such behavior. Panel 2C depicts the average mean single glance durations to the device; the average glance duration for the VAAN is around1.0 seconds, as compared with between about 2.5 seconds and 3.2 seconds for the other systems and comparison tasks. Finally, Panel 2D indicates that, on average, the mean single glance duration to the note card during a destination entry trial with the VAAN was substantially longer than for the other systems or the cellular telephone task.

Lane exceedences represent one measure of degraded vehicle control that may be associated with driver inattention or distraction. Figure 2.3 presents the lane exceedence count averages per trial for the POI destination entry. Panel 3A indicates that age had a significant effect on lane exceedences. Older drivers in the study had, on average, about 8 lane exceedences per 10 trials, as opposed to younger drivers who had a little less than 2 lane exceedences per 10 trials. Panel 3B depicts the average number of lane exceedences per trial as a function of route guidance system device, with 10-digit cellular telephone dialing and manual radio tuning included for comparison purposes. Perhaps the most striking aspect of this panel is that the VAAN was associated with no lane exceedences.

Figure 2.4 shows mean Eyes-off-Road-Time (EORT) results. EORT is the average cumulative length of a trial time spent with the eyes off the road ahead (e.g., looking at the device, note card, etc.). Panel 4A shows older test participants spent about twice as long as younger test participants looking away from the road scene ahead. Panel 4B indicates that, among the route guidance systems, the VAAN was associated with the least amount of EORT, on average, only slightly higher than that for manual 10-digit cellular telephone dialing or radio tuning. Panel 4C again shows the voice

destination entry feature of the VAAN served to minimize the differences between older and younger drivers. Panel 4D presents the average single glance duration to the road scene ahead during the in-vehicle transactions. As can be seen, the VAAN was associated with longer glance durations to the road scene ahead than to any other route guidance system or comparison task. As in-vehicle task demands grow, the driver is often prompted to shorten intermittent glances back to the road scene (perhaps to reduce working memory load), potentially missing safety-relevant objects and events.

2.4 Conclusions

These data suggest voice recognition technology is a viable alternative to visual-manual destination entry while driving. This result is highlighted in test participant subjective assessments that favored voice input over visual-manual methods. However, this study ideally would be replicated and field validated. Further research must also be conducted to examine the effects of voice interaction on the selective withdrawal of attention that degrades object and event detection while leaving visual allocation to the road ahead and vehicle control largely intact. In the interim, these data suggest that destination entry with visual-manual methods is ill-advised while driving.

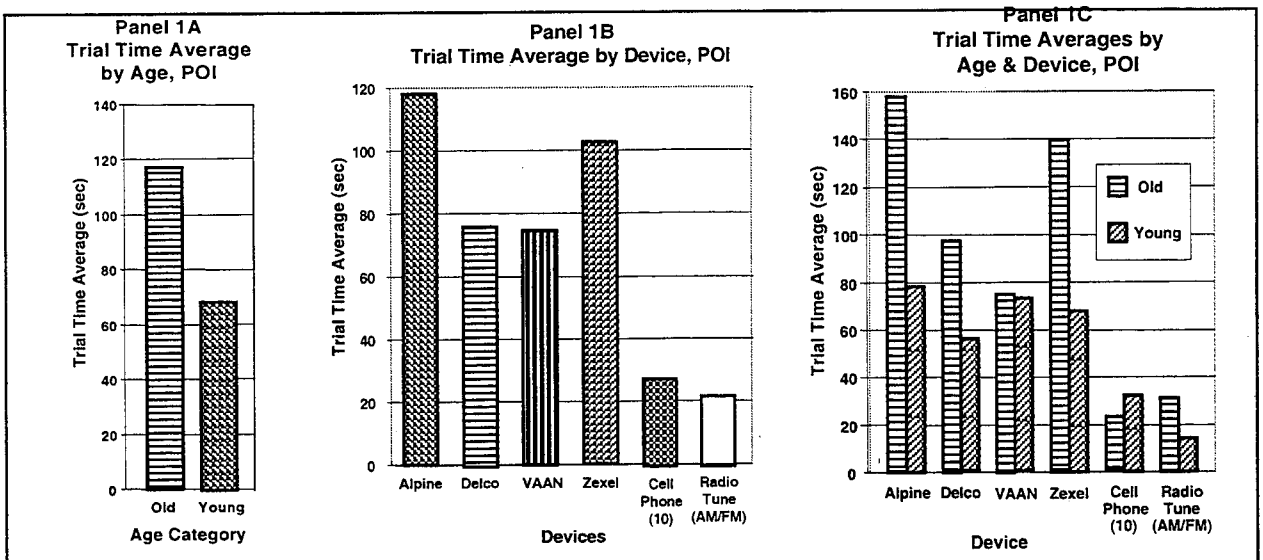


Figure 2.1. Age and Device Effects on Trial Time (i.e., Task Completion Time).

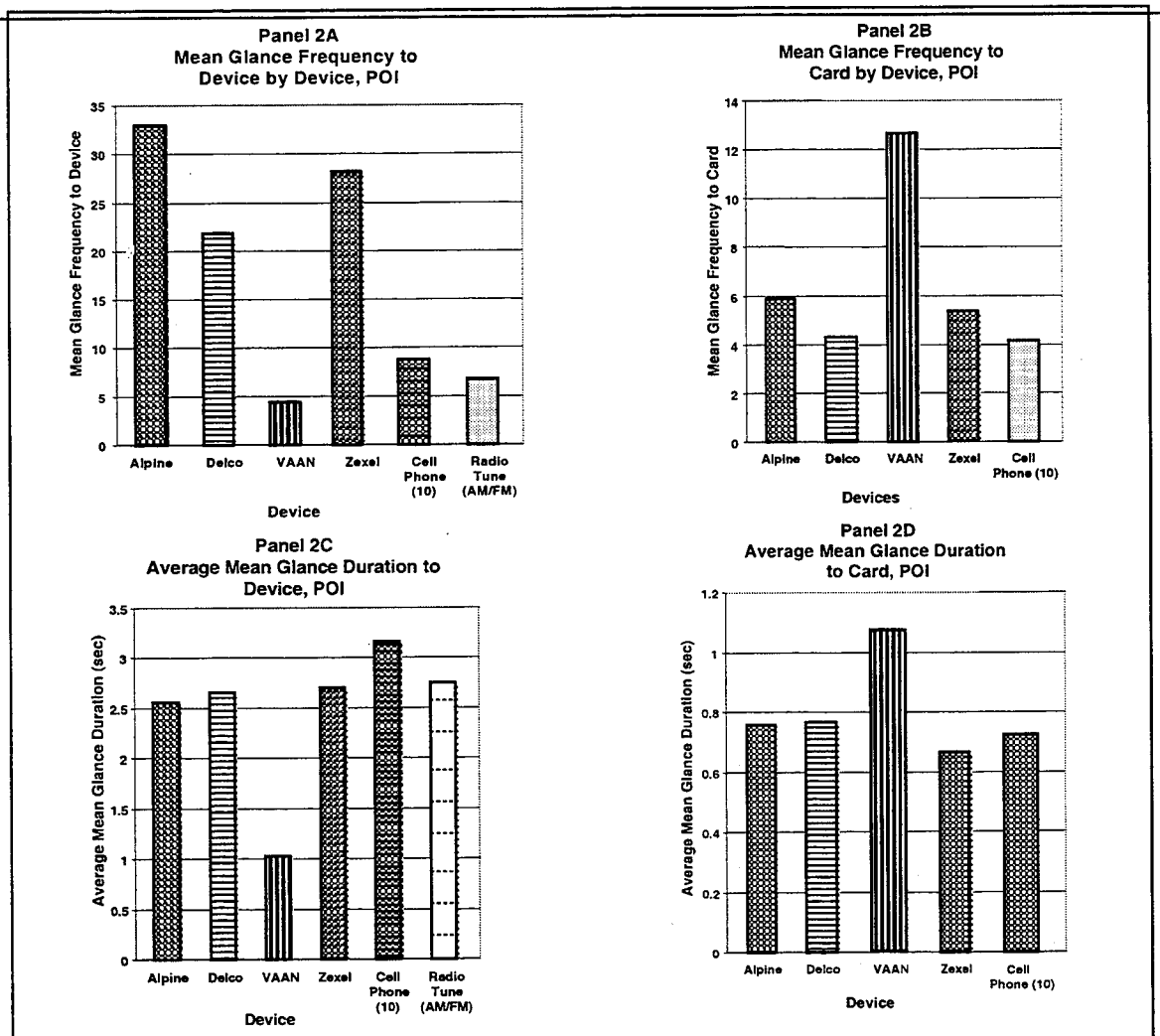


Figure 2.2. Device and Note Card Effects on Glance Frequency and Duration.

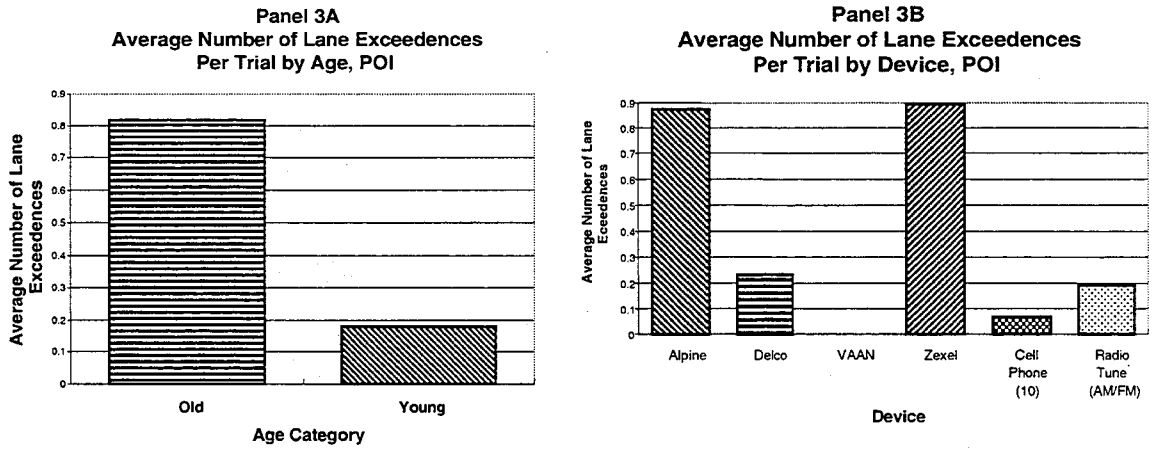


Figure 2.3. Age and Device Effects on Number of Lane Exceedences per Trial.

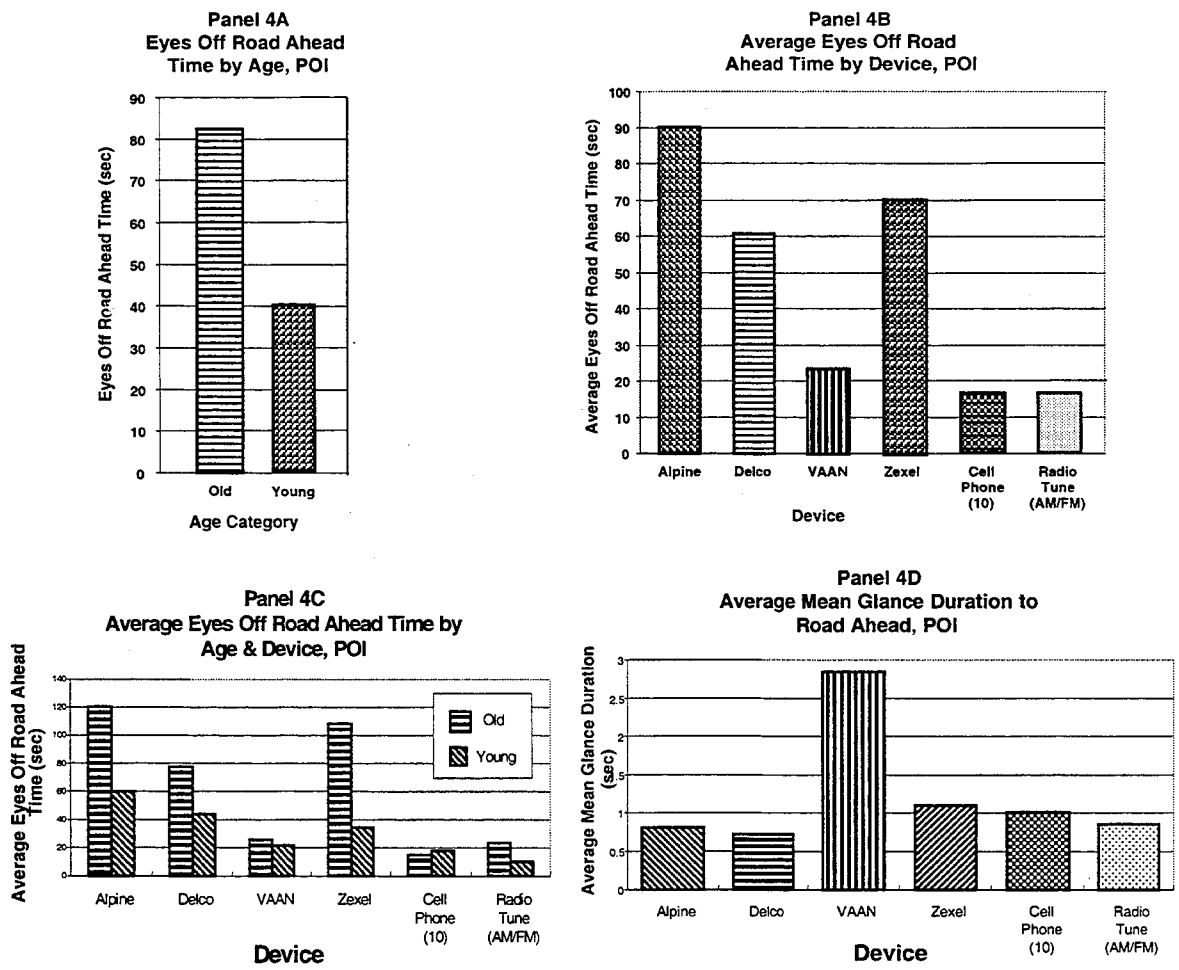


Figure 2.4. Age and Device Effects on Eyes-Off-Road-Ahead Time and Road Glance Duration.

3.0 INDIVIDUAL DIFFERENCES AND IN-VEHICLE DISTRACTION WHILE DRIVING: A TEST TRACK STUDY AND PSYCHOMETRIC EVALUATION

3.1 INTRODUCTION

The proliferation of information and telecommunications systems for use in cars and trucks has made driver distraction a pressing highway safety concern. Driver distraction or workload reflects three major influences: the nature of the in-vehicle device or task, the driving conditions under which that task is pursued, and the individual abilities of the driver. Existing research has focused primarily on assessing driver distraction in terms of the in-vehicle device or task (e.g., Tijerina, Parmer, and Goodman, 1998). Less attention has been devoted to characterizing the driving conditions under which a task might be pursued (e.g., Hulse, Dingus, Fischer, and Weirwille, 1989). Even less research has been conducted to assess how individual differences among drivers might influence their propensity toward distraction.

Research into individual differences and distraction might contribute to highway safety in several ways. For commercial vehicle operations, human abilities tests might be identified that correlate substantially with safe driving. Kahneman, Ben-Ishai, and Notan (1973), for example, found performance on an auditory shadowing task to be significantly correlated with crash involvement among a sample of Israeli truck drivers. Identification of human abilities associated with time-sharing skill in a driving context might lead to new methods of driver training. Individual differences research can also support system design. For example, drivers differ in their spatial abilities as measured by psychometric tests. Such differences manifest themselves when drivers must use moving map displays in route guidance systems. On the other hand, drivers make good use of egocentrically-defined text directions regardless of their spatial abilities (McGehee, personal communication).

Recent research has identified a set of temporal factors in visual perception and cognitive factors might be predictive of real world performances (Kennedy, et al.,1997). It was desired to determine the extent to which such tests, provided in a computerized battery termed PATSYS, might be indicative of performance during in-vehicle device use while driving.

3.2 METHOD

3.2.1 Test Participants: Sixteen (16) TRC test drivers participated. These drivers were hourly employees with valid driver's licences and generally less than 2 years of TRC driving experience. There were equal numbers of males and females in each of two age categories: Younger (35 years or younger) and Older (55 years or older). None of the test participants owned or had significant prior experience with route guidance systems or cellular telephones prior to this study.

3.2.2 Test Vehicle and Instrumentation: The test vehicle was a 1993 Toyota Camry, equipped with MicroDAS instrumentation (Barickman, 1998) which captured travel speed, lane position, and lane exceedences, as well as video of the road scene and driver eye glance behavior at a sampling rate of 30 Hz. Eye glance video was later manually reduced.

3.2.3 Route Guidance Systems: Four (4) unmodified, commercially available route guidance systems. The dash mounted Delco Telepath 100® consisted of a 3-line LCD display to present menu items, scrolled by means of a bezel-mounted rotary knob and selected by pressing an Enter key. The Alpine NVA-N751A® incorporated a free-mounted 5.6 inch active matrix color display without bezel keys. It displayed an alphanumeric keyboard and entries were made by scrolling from key to key with a joystick mounted on a remote control unit; pressing down on the joystick registered a character or selection. If sufficient alphanumerics were entered for the system to estimate candidate destinations, these were presented as an alphabetized scrolling list of 3 items at the bottom of the display of the alphanumeric keyboard screen. The Zexel Navmate® consisted of a free-mounted 4 inch diagonal full color LCD screen with a set of bezel control keys, including a central "left, right, up, down" key and an Enter key. Both the Zexel and Alpine systems were mounted on a gooseneck pedestal bolted to the floor board between the driver and passenger. The Zexel system presents menu options for destination entry type and city, followed by a scrolling display of numerically and alphabetically arranged destinations generally presented 11 to13 lines at a time. The driver presses the Enter key to make a selection. Finally, the dash mounted Clarion Eclipse® Voice Activated Audio Navigation (VAAN) system used voice recognition and output exclusively; there was no visual display. Keywords would activate the VAAN for destination entry. Destinations were entered by spelling them. The VAAN emphasized precise spelling of a destination; each letter uttered by the driver would be proceeded by a beep to acknowledge receipt of the input. The driver uttered "verify" to conclude an entry. The system would eventuate in a spoken list of best-guess candidate destinations for selection by the driver via YES or NO verbal responses.

The last three of these systems allowed for entry of a street address, intersection, or point of interest (attraction, restaurant, hotel, etc.). Thus, three types of tasks (address, intersection, point of interest) were included as suitable for comparisons among the systems. The Delco system only supported point of interest selection. Also, two additional tasks were included for comparison purposes: tuning a radio to a specific band and frequency using a modern "Seek" function on the Clarion Eclipse system; and manually dialing a cellular telephone (a 10-digit number on a handwritten note card) using a cordless AUDIOVOX Model MVX-500.

The PATSYS test battery was used to conduct the psychometric evaluation (See Table 3-1). This battery was run on a Gateway 2000 E-3110 personal computer with a Vivitron 17-inch diagonal high resolution color monitor. For further details of these tests, see Kennedy, Silver, and Ritter (1995), Turnage and Kennedy (1995), and Kennedy, Turnage, and Lane (1997).

Table 3-1. PATSYS Test Battery: Temporal and Cognitive Subtests.

TEST NAME	TEST DESCRIPTION	SCORING
Dynamic Visual Acuity (DVA):	This test varied the presentation time between the letter "C" presented on the left of the video screen and a letter "C" presented on the right of the video screen. The participant's task was to determine if the C's were facing in the same or opposite direction. DVA refers to the ability of an observer to resolve fine detail in an object when there is relative motion between them. This may be important as a driver repeatedly shifts gaze from the road scene to inside the vehicle.	The participant's score was the fastest presentation time for correct responses. A lower score indicates better dynamic visual acuity.
Simultaneity (SIMU):	This test presented two open boxes 33 mm apart which were alternately flashed on the screen for 60 msec. The interstimulus interval (ISI) for onset of the two boxes was manipulated. The basis for this test is that relatively large temporal differences are needed before an observer can reliably perceive two stimuli as non-simultaneous. Reflects visual processing speed, acuity.	The participant's score was the lowest time value or ISI when the two boxes appeared to be on simultaneously. A lower score indicates better temporal acuity.
Bistable Stroboscopic Motion (STR)	This test presented an array of boxes that were alternately cycled. Frame one consisted of three horizontal elements of boxes with equal center-to-center distances. Frame two had identical elements had identical elements shifted to the right by a distance equal to the center-to-center separation between stimuli. Participants responded by keyboard presses whether they perceived "element" motion (appearing as four boxes) or "group" motion (appearing as a set of three boxes that alternatively shifted back and forth laterally one box-width).	The point of transition from one type of motion to the other was collected as a threshold value. A lower score represents greater sensitivity.
Phi Phenomenon (PHI)	Square boxes, 33 mm apart on the video display were presented to the left and right of a fixation point. Through a set of response keys, the participant would adjust the interstimulus interval (ISI) to the point where the boxes appeared to transition from moving successively to a single box moving back and forth. Reflects visual processing speed.	The value of the ISI at the fifth reversal was the participant's score. A lower score would signify better temporal acuity.
Masking Test (MSK)	In this test, two vertical lines .075" in length and 0.05" in width were presented. A horizontal line 0.05" in length extended from the midpoint of either the left or right vertical line. After a brief period, the lines were replaced by a complex pattern of dots (the mask). The screen went blank and the participant was instructed to press the left or right arrow keys depending on whether the horizontal line was on the left or right vertical line. Masking is the interference in the perception of one briefly presented stimulus by a second, succeeding stimulus briefly presented nearby in time and space. This may be important as the driver attempts to retain information in working memory while glancing between road scene and in-vehicle device.	The ISI between the target and the mask was varied and the participant's score was the lowest ISI for correct responses. A lower score signified better temporal acuity.
Grammatical Reasoning (GR)	This test employs five grammatical transformations on statements about the relationship between two letters "A" and B" For example, There are 32 possible items arranged in random order. The participant assesses the correctness of the statement by pressing the "T" key for true statements or the "F" key for false statements. This measures higher cognitive processes of deductive reasoning. It may reflect a general effect of greater cognitive capability on task completion.	The participant's score is the number correct out of 32 statements. A higher score is indicative of greater capacity of higher cognitive processing.
Mannikin Test (MNK)	A simulated human figure (a sailor) is presented in either full-front or full-back orientation on the screen. The figure is shown holding three hearts, diamonds, clubs, or spades, different patterns in each raised hand. One of the two patterns held matches a pattern which appears on a podium the figure stands on. The participant indicates which hand is holding the pattern matching that in the podium by pressing the appropriate key. This test appears to measure ability in mental rotation and related transformations. MNK scores might reflect the ability of a driver to reorient spatially between glances to an in-vehicle display layout and the road scene.	The participant's score is the number correct out of 16 trials. A higher score signifies better spatial ability.
4-Choice Reaction Time (RT4)	On this test, four outlined boxes are displayed above the numbers 1, 2, 3, and 4. At random intervals, one of the boxes illuminates, i.e. changes from outline to filled. The participant presses a corresponding key as quickly and accurately as possible. This test assesses the participant's speed of information processing to make a response from multiple alternative stimuli, depending on which alternative is signaled. Speed of cognitive processing is ubiquitous as an contributor to cognitive task performance.	The participant's score is the latency between when a box illuminates on the screen and when the corresponding key is pressed on the keyboard. Shorter reaction times generally represent faster processing

Pattern Comparison (PC)	In this test, a pair of eight-dot patterns are presented and the test participant indicates on the keyboard whether the two patterns are the same or are different. This is a test of perceptual speed. This may be important as driver's compare entry data to device display feedback.	The participant's score is the number of pairs correctly identified as similar or different. Higher scores imply greater perceptual speed.
Code Substitution (CS)	This test involves a display of nine characters on the top of the screen and beneath them the numbers 1 through 9 in parentheses. Under the code are two rows of characters with empty parentheses beneath them. The participant inserts the number associated with the character from the code displayed at the top of the screen. This test appears to assess working memory and perceptual speed. It may reflect the capability of a driver to keep track of task components.	The score is made up of the number of correctly matched digits to their corresponding letters. Higher scores imply greater perceptual speed and working memory.

3.2.4 Test Track: The TRC 7.5-mile multi-lane test track is in the form of an oval with banked curves at either end and with unbanked straightaways that measure approximately 2.0 miles each. The test track is comprised of three 12-ft wide concrete lanes with a fourth inner blacktop lane for use in the event of vehicle breakdowns or required stops. The test vehicle for this study operated in lane 1 (adjacent to the innermost blacktop lane) and changed lanes only as needed for normal track operations and safety. The test participant was asked to drive at approximately 45 mph on the straightaways and accelerate to 60 mph on the curves, provided that any requested tasks are completed by the time the test vehicle enters a curve. Otherwise, the driver was to maintain 45 mph and attempt to complete the requested in-vehicle task. Traffic density tended to be light relative to open road driving. However, travel speeds for other vehicles of the track might vary greatly, vehicles involved with other testing could slow, stop, or move to the blacktop lane abruptly, and track repair and roadside obstructions had to be avoided. Faster traffic drove on the outer lanes of the oval. Data collection was scheduled for between 8:00 am and 4:30pm weekdays.

3.2.5 Procedure: Prior to the data collection runs, the test participant signed an informed consent form (see Appendix A) and the experimenter familiarized the test participant with each navigation system. Each test participant then completed 12 practice data entry tasks per system (four for each destination category), entered while the vehicle was parked. This training was done in two phases (morning and afternoon); so, two systems were reviewed prior to each half of the test track trials. On the 7.5 mile track, the order of trials were counterbalanced across the four route guidance systems (Zexel, Alpine, Delco, and VAAN), destination entry category (point of interest, intersection, and street name targets), and target (Target A or Target B within a category). All trials with a given system were executed before moving on to another system; the destination type and targets within destination type were counterbalanced to control for order effects. The cellular phone and radio tuning tasks were interspersed between destination entry trials on an opportunistic basis by the experimenter in a quasi-random fashion. Prior to leaving for the test track, the destinations were presented to the test participant in 18-point Times Roman font and the test participant was asked to write in his or her own hand each destination on a separate index card, as well as the 10-digit unfamiliar telephone number, such that they would be able to read from it while driving. A task began when the ride-along experimenter gave the driver a hand-written card or a radio tuning task was requested orally by the ride-along experimenter. The task ended when the request had been fulfilled, as indicated by an event marker triggered by the experimenter. Requests for tasks were generally made when the test participant was exiting a curve onto a straightaway segment of the test track. After test track data collection was completed, the test participant answered the subjective assessment questions and was released. Each test participant was invited back subsequent to the test track trials and administered the battery of temporal acuity and cognitive tests. The battery of tests

was administered four times in a single day. The results from the last of the four rounds of testing were used for data analysis.

3.2.6 Test Track Measures, Test Battery Measures: Four response measures from the test track study were selected for analysis: in-vehicle task completion time per trial (TASKTIME, seconds), mean average glance duration to the device during task completion (MNGLNCTM, seconds), glance frequency or number of glances to a given in-vehicle device per trial (GLNCFREQ), and number of lane exceedences or departures per trial or task completion (NEXCEED). The previously mentioned test battery subtests generated latencies for all TEMPORAL tests (see Kennedy, et al., 1995 for explanation of these) and for the RT4 test; all other cognitive tests were scored in terms of number of trials correct over a fixed period of testing (not a fixed number of trials).

3.3 RESULTS

The data were analyzed in terms of correlation and regression. Table 3-2 shows the matrix of intercorrelations among test track and test battery measures. This table reveals that among the test track measures a) task time is highly correlated with glance frequency to the device, b) both are moderately correlated with number of lane exceedences (NEXCEED), and c) mean glance duration is not correlated with any of the other test track measures. This pattern of results is comparable and consistent with other studies examining the effects of in-vehicle device use while driving (Green, 1998).

Among the battery of temporal and cognitive tests there is a moderate degree of intercorrelation among the temporal tests and high intercorrelations among the cognitive tests. However, the intercorrelations across the two subsets of tests are generally lower than within each subset, indicating that they reflect distinct aspects of human performance.

An all-possible-regressions analysis was carried out using the PROC REG procedures in SAS. For each of the test track measures, all possible models were assessed for the various combinations of those test battery tests with a statistically significant correlation ($\alpha \leq 0.05$). The criterion used to select the "best" model was the adjusted R^2 criterion. The adjusted R^2 criterion is equivalent to finding the set of predictor variables that minimizes the residual mean square error for the model (Montgomery and Peck, 1992).

In Table 3-2 reasonable patterns of correlation appear with task completion time. TASKTIME worsens (i.e., increases) as DVA, STR, MSK, and RT4 scores worsen (i.e, increase) and TASKTIME improves (i.e., decreases as CS, PC, GR, and MNK scores improve (i.e., increase). The only anomaly is PHI; TASKTIME decreases as PHI scores worsen (i.e., increase). Overall, as measures of temporal acuity, perceptual speed, working memory, speed of processing, spatial abilities, and higher cognitive processes improve, TASKTIME decreases reliably. However, the proportion in TASKTIME variability that covaries with a set of such regressors is modest. The "best" subset of regressors for TASKTIME were PHI, RT4, and GR with Multiple $R = 0.35$, Adjusted $R^2 = 0.12$.

Correlations between significant PATSYS tests and glance frequency (GLNCFREQ) also are consistent with intuition. Glance frequency worsens (i.e., increases) as DVA, STR, MSK, and RT4 scores worsen while GLNCFREQ improves (i.e., decreases) as CS, PC, GR, and MNK scores

improve. However, the proportion of variability in GLNCEFREQ that covaries with the "best" regression model is only about 10 percent. The "best" set of predictors for GLNCFREQ were STR MSK, and PC, with Multiple R = 0.33, Adjusted $R^2 = 0.10$.

The only variable selected as a regressor for mean glance time (MNGLNCTM) was MSK, with $R^2 = 0.03$. As masking scores worsened (i.e., increased), so did mean glance time.

Table 3-2. Intercorrelation Matrix for Test Track and Test Battery Measures.

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0

	TASK TIME	GLNC FREQ	NEX CEED*	MNGL NCTM	DVA	SIMU	PHI	STR	MSK	CS	PC	RT4	GR	MNK
TASK TIME	1.0000 0.0	0.8587 0.0001	0.3797 0.0001	-0.0485 0.3533	0.1995 0.0001	-0.0177 0.7346	-0.1420 0.0064	0.1766 0.0007	0.2241 0.0001	-0.2823 0.0001	-0.3021 0.0001	0.3055 0.0001	-0.2609 0.0001	-0.2932 0.0001
GLNC FREQ		1.0000 0.0	0.2972 0.0001	-0.0688 0.1879	0.1975 0.0001	-0.0173 0.7413	-0.0924 0.0767	0.2420 0.0001	0.1319 0.0113	-0.2358 0.0001	-0.2836 0.0001	0.2699 0.0001	-0.2629 0.0001	-0.2129 0.0001
NEX CEED*			1.0000 0.0	0.0092 0.8630	0.2548 0.0001	-0.0526 0.3242	-0.0759 0.1545	0.1086 0.0414	0.4128 0.0001	-0.2266 0.0001	-0.1764 0.0009	0.2654 0.0001	-0.0601 0.2598	-0.2966 0.0001
MNGL NCTM				1.0000 0.0	0.1349 0.0096	-0.0317 0.5447	-0.0359 0.4919	0.0446 0.3936	0.1731 0.0009	-0.0780 0.1355	-0.0301 0.5649	0.069 0.1842	0.0201 0.7004	-0.1025 0.0494
DVA					1.0000 0.0	-0.0995 0.0564	-0.1633 0.0017	0.4981 0.0001	0.5762 0.0001	-0.4655 0.0001	-0.4251 0.0001	0.4669 0.0001	-0.2753 0.0001	-0.3827 0.0001
SIMU						1.0000 0.0	0.0533 0.3083	-0.0696 0.1824	-0.1565 0.0026	0.0371 0.4779	-0.0584 0.2641	0.2107 0.0001	-0.0130 0.8029	0.2262 0.0001
PHI							1.0000 0.0	-0.0618 0.2367	-0.2217 0.0 001	0.1981 0.0001	0.2448 0.0001	-0.2004 0.0001	-0.0977 0.0611	0.2700 0.00 01
STR								1.0000 0.0	0.4245 0.0001	-0.2676 0.0001	-0.4204 0.0001	0.6460 0.0001	-0.3600 0.0001	-0.1956 0.0002
MSK									1.0000 0.0	-0.6591 0.0001	-0.6177 0.0001	0.6849 0.0001	-0.2829 0.0001	-0.6757 0.0001
CS										1.0000 0.0	0.8497 0.0001	-0.5949 0.0001	0.6392 0.0001	0.7993 0.0001
PC											1.0000 0.0	-0.7535 0.0001	0.6474 0.0001	0.7873 0.0001
RT4												1.0000 0.0	-0.4802 0.0001	-0.6480 0.0001
GR													1.0000 0.0	0.5884 0.0001
MNK														1.0000 0.0

* NOTE: Number of Observations = 368 except NEXCEED = 353.

TASKTIME: In-Vehicle Task Completion Time; GLNCFREQ: Number of glances to device to complete task; NEXCEED: Number of lane exceedences during task completion; MNGLNCTM: Mean glance time to device during task completion.

Correlations between significantly correlated PATSYS tests and the incidence of lane exceedences (NEXCEED) were all sensible in sign. NEXCEED measures worsened (i.e., exceedences increased) as DVA, STR, MSK, and RT4 scores worsened (i.e., increased). NEXCEED measures improved (i.e., decreased) as CS, PC, and MNK scores improved. The best subset of regressors identified were MSK, PC, and MNK, with a multiple R = 0.44, adjusted $R^2 = 0.19$.

3.4 DISCUSSION

This study represents an attempt to assess the explanatory power of individual differences in both temporal acuity and cognitive abilities in terms of various measures of driver distraction or workload while using a variety of in-vehicle devices. The variability shared in common between a given measure of test track performance and the "best" subset of test battery measures is modest at best. This perhaps reflects the relative contribution of individual differences (as measured by these tests) to in-vehicle task completion while driving. This finding is consistent with other research into individual differences and highway safety (Elander, West, and French,1993). It would not be surprising to find that the specifics of the task and driving conditions at the time of task execution, combined with driver motivation, fatigue, and the like command a much larger share of the variability in task outcomes. There is also random errors that arise in device use and a variation in error recovery that also increase response variability.

When each dependent measure was examined within the context of specific test battery components, there was high face validity to predictor sets. Thus, better task time was associated with better temporal acuity, faster processing and higher cognitive capabilities. Likewise, reduced glance frequency was associated with better dynamic visual and temporal acuity, better pattern comparison performance and faster processing of information. These relationships and degree of overlap suggests that with greater refinement, efficiency and packaging of the test battery it may be possible to tune in-vehicle tasks to the specific cognitive and temporal capabilities of individual drivers, a step towards building truly "intelligent" systems. Future work should examine such refinements and explore the more subtle relationships between specific task demands and predictor sets.

4.0 Preliminary Evaluation of the Proposed SAE J2364 15-Second Rule for Accessibility of Route Navigation System Functions While Driving

4.1 Introduction

New and increasingly powerful information and telecommunications technologies are finding their way into cars and trucks. These new technologies include route navigation systems. Such systems accept destinations entered by a driver, generate a route based on some prescribed criteria (e.g., minimize travel time, minimize distance), and provide the driver with turn-by-turn directions through various types of visual or auditory displays (or both). Some systems integrate up-to-date traffic advisories into their route guidance and provide yellow pages information on selected points of interest. On the positive side, these systems can be very useful to a traveler in unfamiliar territory or coping with changes in traffic flow. On the negative side, interaction with such systems may pose a significant distraction to the driver. From a safety standpoint, it is important to have methods to assess the distraction potential a particular system design poses to a driver. Such methods can provide useful guidance on system design or use to reduce the workload imposed on the driver.

It was this need for evaluation methods and criteria for route guidance systems that prompted the Society of Automotive Engineers (SAE) Safety and Human Factors Committee to initiate a project to develop a standard or recommended practice to determine whether or not a particular route guidance system function should be accessible to the driver while driving. A subcommittee of representatives from the academic, automotive, and navigation technologies sectors was convened and a consultant was retained to develop the standard. The result is draft standard SAE J2364 (Green, 1999a; Society of Automotive Engineers, 1999).

Draft standard SAE J2364 is intended for use to evaluate if a given function of a particular route navigation system should or should not be accessible while driving. SAE J2364 proposes that if an in-vehicle task, executed <u>without</u> concurrent driving (e.g., in a parked vehicle), can be completed in 15 seconds or less by a sample of drivers, then that in-vehicle task may be accessible while driving (Green, 1999b). This rule is based on a review of published literature, a reanalysis of published data sets, and expert judgements of the SAE Safety and Human Factors subcommittee tasked to develop the standard. A key feature of the proposed "15-second rule" is its ease of implementation as an evaluation approach (Green, 1999a).

The decision to examine single task completion time (hereafter referred to as static completion time) rather than task completion time while timesharing with the driving task (hereafter referred to as dynamic completion time) was driven in part by a desire to support GOMS modeling. GOMS stands for Goals, Operators, Means, Selection rules and the basic methodology was originally developed by Card, Moran, and Newell (1983). GOMS models are developed from a detailed task analysis of a task or transaction. A database of component task times is then used with the task analysis information to generate predictions of total completion times or statistical predictions. In this sense, GOMS modeling is a cognitive extension of more traditional predetermined time systems methods applied in industrial time-and-motion analysis (Smith, 1978). GOMS modeling requires for its implementation a) a focus on single task performance rather than performance while timesharing with another task, b) errorless performance (GOMS modeling does not handle error), and c) skilled

performance in the sense of a known strategy for task completion. While these requirements may be seen as limitations, GOMS modeling provides quantitative design guidance early on in the product development cycle when an operational prototype may not yet exist. A companion standard to J2364 has been drafted which outlines the GOMS modeling approach in the context of the 15-second rule. This is draft SAE standard J2365 (Society of Automotive Engineers, 1998a). Green (1999b) has provided a worked example of GOMS modeling of destination entry into a route guidance system, including some preliminary task times suitable for the automotive environment..

While application of the GOMS model simplifies device function or task evaluation, there is a paucity of data available on the relationship between static completion time for tasks and performance while driving. The original version of SAE J2364 prescribed a method of assessment that was characterized as a "reasonable worst case" (SAE, 1998b). Specifically, the draft standard at the end of 1998 indicated that a sample of 10 test participants between the ages of 55 and 65 years would complete the target tasks after device orientation and 5 practice trials and their static completion times on a subsequent trial would be used as the data upon which to recommend whether or not a task should be accessible to the driver while the vehicle is in motion. This approach was empirically tested for the first time in the study reported below.

4.2 Approach

4.2.1 Test Participants: Ten (10) test participants were recruited from the TRC pool of entry-level test drivers. Five females (ages 55, 56, 58, 62, and 63 years) and five males (56, 58, 64, 65, and 69 years) were selected for participation. These drivers were hourly employees with valid driver's licences and generally less than 2 years of TRC driving experience. None of the test participants owned or had significant prior experience with route guidance systems prior to this study.

4.2.2 Test Vehicles and Test Track: The subject vehicle was a 1993 4-door Toyota Camry operated with cruise control disabled. In addition, two confederate vehicles were also used during test track testing. A 1991 Acura Legend driven by a member of the research team served as a lead vehicle preceding the subject vehicle. A 1996 Ford Taurus was used as a following vehicle trailing after the subject vehicle. The purpose of the lead vehicle was to provide some measure of driving task load to the test participant. The following vehicle was used for an observer to manually count lane exceedences or lane departures that occurred during the execution of a task.

Testing was again conducted on the TRC 7.5-mile track. The test vehicles for this study operated in lane 1 (adjacent to the innermost blacktop lane) and changed lanes only as needed for normal track operations and safety. The test participant was asked to drive at approximately 45 mph on the straightaways and to accelerate to 60 mph on the curves, provided that any requested tasks were completed by the time the test vehicle enters a curve. Otherwise, the driver was to maintain 45 mph, maintain a self-determined "safe" following distance from the lead vehicle, and attempt to complete the requested in-vehicle task. Traffic density of non-confederate vehicles on the TRC track tended to be light relative to open road driving. However, travel speeds for other vehicles of the track might vary greatly, vehicles involved with other testing could slow, stop, or move to the blacktop lane abruptly, and track repair and roadside obstructions had to be avoided. Faster traffic drove on the

outer lanes of the track. Data collection was scheduled for between 8:00 am and 4:30pm weekdays in February, 1999. Dry pavement conditions were required for data collection.

4.2.3 Route Guidance Systems and Other In-Vehicle Tasks: Four (4) unmodified, commercially available route guidance systems, each with a different destination entry and retrieval logic and driver interface, were used in the test. The dash mounted Delco Telepath 100® consisted of a 3-line LCD display to present menu items, scrolled by means of a bezel-mounted rotary knob and selected by pressing an Enter key. The Alpine NVA-N751A® incorporated a free-mounted 5.6-inch active matrix color display without bezel keys. It displayed an alphanumeric keyboard and entries were made by scrolling from key to key with a joystick mounted on a remote control unit; pressing down on the joystick registered a character or selection. If sufficient alphanumerics were entered for the system to estimate candidate destinations, these were presented as an alphabetized scrolling list of 3 items at the bottom of the display of the alphanumeric keyboard screen. The Zexel Navmate® consisted of a free-mounted 4-inch diagonal full color LCD screen with a set of bezel control keys, including a central "left, right, up, down" key and an Enter key. Both the Zexel and Alpine systems were mounted on a gooseneck pedestal bolted to the floor board between the driver and passenger. The Zexel system presented menu options for destination entry type (described below) and city, followed by a scrolling display of numerically and alphabetically arranged destinations generally presented 11 to13 lines at a time. The driver presses the Enter key to make a selection. Finally, the dash mounted Clarion Eclipse® Voice Activated Audio Navigation (VAAN) system used voice recognition and output exclusively; there was no visual display. Keywords would activate the VAAN for destination entry. Destinations were entered by spelling them. The VAAN emphasized precise spelling of a destination; each letter uttered by the driver would be proceeded by a beep to acknowledge receipt of the input. The driver uttered "verify" to conclude an entry. The system would eventuate in a spoken list of best-guess candidate destinations for selection by the driver via YES or NO verbal responses.

The last three systems allowed for entry of a street address, intersection, or point of interest (attraction, restaurant, hotel, etc.). Thus three types of entry tasks (address, intersection, point of interest) were included as suitable for comparison among the systems. The Delco system only supported point of interest selection.

Five additional tasks were included for comparison purposes. These were:

1. Manually tuning a radio to a specific AM band frequency. This was accomplished using a modern "Seek" function that is part of the radio portion of the Clarion Eclipse system. The Clarion Eclipse system was mounted low in the center console area.
2. Manually tuning a radio to a specific FM band frequency again using the radio portion of the Clarion Eclipse system.
3. Manually dialing an unfamiliar, 10-digit, phone number which had been handwritten on a note card using a cordless AUDIOVOX Model MVX-500 phone.
4. Manually dialing familiar, 7-digit, phone number (e.g., a home phone number) again using the AUDIOVOX phone.
5. Manipulate the heating, ventilation, and air-conditioning (HVAC). This task involved changing the fan speed, recirculation toggle, and defroster.

In all, fifteen (15) different in-vehicle tasks were used for this study.

4.2.4 Response Measures: For this study, three different response measures were taken. Static Completion Time was defined as the time taken while performing a task in a parked vehicle. The start of the task was operationally defined as the point in time where the experimenter either handed the test participant a hand-written card with the appropriate information to be entered or verbally requested the HVAC adjustments or dialing home task. The end of the task was determined by the experimenter when the goal had been reached. Dynamic Completion Time was defined as the time taken, while performing an in-vehicle task and concurrently driving on the 7.5-mile test track, from the start of a task until the completion of that task as previously defined. Both of these measures were collected by the experimenter with a digital stop watch. The third response measure of interest was Lane Exceedence count per trial, i.e., the number of times the subject vehicle crossed either lane line during completion of a in-vehicle task while concurrently driving. This measure was recorded by a member of the research team observing the subject vehicle from the following confederate vehicle.

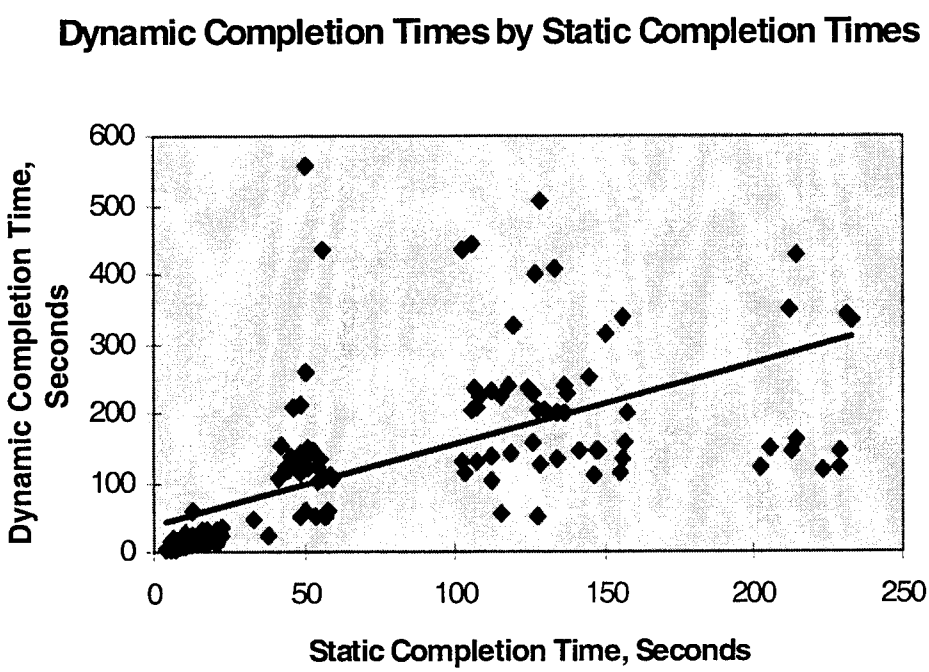
4.2.5 Procedure: Prior to the data collection runs, the test participant signed an informed consent form (see Appendix B) and the experimenter familiarized the test participant with each in-vehicle system. This included a structured walk-through on the procedures associated with each task, demonstration of the task by the experimenter, and opportunity for questions and answers. Each test participant then completed 5 practice trials per task per system, completed while the vehicle was parked in a car bay at TRC. This training was done in two phases (morning and afternoon); so two route guidance systems were reviewed prior to each half of the test track trials. On the 7.5 mile track, the order of trials were counterbalanced across the four route guidance systems (Zexel, Alpine, Delco, and VAAN) and the destination entry category (point of interest, intersection, and street name targets). All trials with a given system were executed before moving on to another system; the destination type and targets within destination type were counterbalanced to control for order effects. The cellular phone, radio tuning, and HVAC tasks were interspersed between destination entry trials on an opportunistic basis by the experimenter in a quasi-random fashion. Prior to leaving for the test track, the destinations were presented to the test participant in 18-point Times Roman font and the test participant was asked to write in his or her own hand each destination on a separate index card, as well as the 10-digit unfamiliar telephone number, such that they would be able to read from it while driving. A destination entry task began when the ride-along experimenter gave the driver a hand-written card containing a destination. A radio tuning task began when the experimenter handed the test participant a note card indicating the band (AM or FM) and radio frequency to be dialed. An unfamiliar 10 digit dialing task began when the experimenter handed the test participant a note card with the 10-digit unfamiliar number to be dialed on the cellular telephone. The experimenter orally requested the HVAC adjustment, or 'dial home' task. A task ended when the request had been fulfilled, as determined by the experimenter. Requests for tasks were generally made when the test participant was exiting a curve onto a straightaway segment of the test track. After test track data collection was completed, the test participant completed a debriefing and was released.

4.2.6 Data analysis: Regression analysis was carried out on each pair of the three response measures. Additionally, a signal detection measure of J2364's diagnostic sensitivity was carried out.

4.3 RESULTS

Correlation and regression analyses were carried out on each pair of the response variables. Prior to these analyses, the data were scanned for outliers or anomalous trials. One test participant had great difficulty being understood by the VAAN voice system; hence no completion time data were available for that person's VAAN tasks. Furthermore, the authors made the decision to exclude trials that involved static completion times greater than 240 seconds or trials that involved more than 10 lane exceedences on the test track. This was done on the grounds that such trials represented egregiously poor performance. Note, however, that analyses carried out on the full data set (i.e., with outliers) do not appear to contradict or alter the fundamental findings or conclusions to be drawn here.

4.3.1 The Relationship between Static Completion Time and Dynamic Completion Time: Figure 4.1 depicts the scatter plot and simple linear regression line for Dynamic Completion Time as a function of Static Completion Time for the same task. As indicated, a simple linear regression model using static completion time to predict dynamic completion time had an $R^2 = 0.39$. Thus, the correlation between static completion time and dynamic completion time is positive, as might be expected. However, there is a substantial standard error. Furthermore, in some instances the dynamic completion times were shorter than the static completion times. Reasons for this will be addressed in the Discussion section.



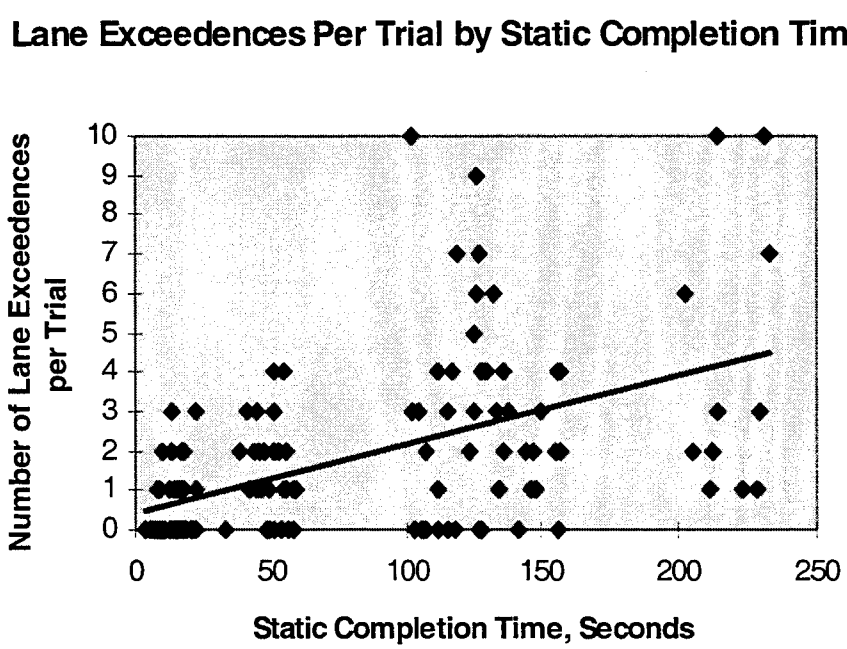
Est. Dynamic Time = 37.56 + 1.16*Static Time, R^2 = 0.39, se = 95.22

Figure 4.1. Regression and Scatter Plot of Dynamic Completion Time vs. Static Completion Time.

4.3.2 The Relationship between Static Completion Time and Lane Exceedences: Figure 4.2 shows the scatter plot and best-fitting linear regression of number of lane exceedences per trial as a function of static completion time. A linear regression model relating number of lane exceedences per trial as a function of static completion time had an $R^2 = 0.27$. This suggests that almost three quarters of the variability in lane departures remains unaccounted for by knowing static completion time. Again there is a large standard error of estimate.

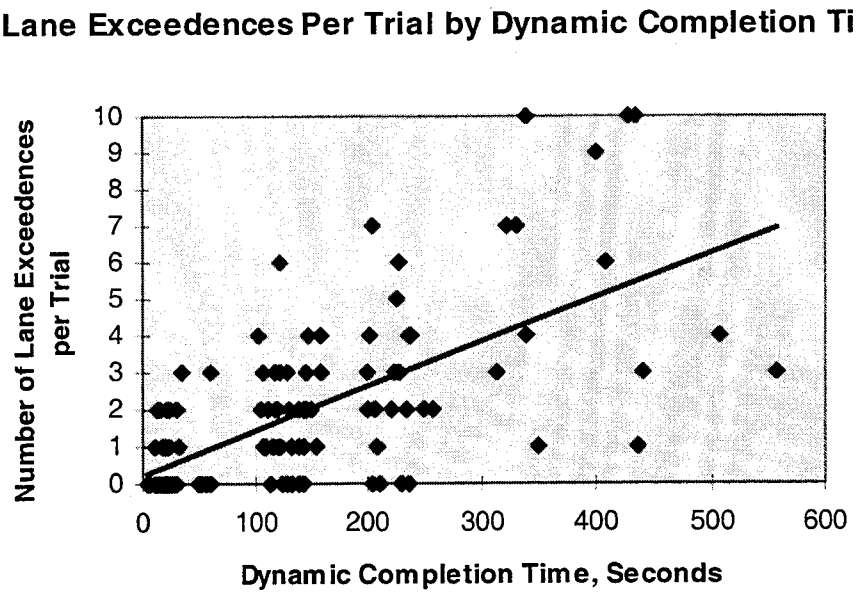
4.3.3 The Relationship between Dynamic Completion Time and Lane Exceedences: Figure 4.3 shows the scatter plot and line of best fit for number of lane exceedences per trial as a function of dynamic completion time. A linear regression model relating number of lane exceedences per trial as a function of dynamic completion time had an $R^2 = 0.43$. Use of dynamic completion time as a predictor improves the percentage of response variability accounted for when compared to static completion time. It nonetheless remains less than 50%.

4.3.4 Signal Detection Analysis of the 15-Second Rule: The Theory of Signal Detection (TSD) (McNicol, 1972) provides quantitative methods for modeling diagnostic system sensitivity and decision bias. Sensitivity refers to the diagnostic system's ability to detect a signal from a background of noise. Decision bias refers to the system's willingness to declare "signal" on a given trial or sample. Originally developed to assess radar systems, TSD is a generally applicable and mathematically rigorous modeling system that has been successfully applied to hardware, software, and human performance, singly and in combination in many real world contexts (Swets, 1996). A brief introduction to the theory and methods of TSD are provided before moving to the application to the 15-second rule.



Est. Lane Exceedences = 0.417 + 0.017*Static Time, R^2 = 0.27, se = 1.88

Figure 4.2. Regression and Scatter Plot of Lane Exceedences per Trial vs. Static Completion Time.



Est. Lane Departures = 0.243 + 0.012*Dynamic Time, R^2 = 0.43, se = 1.65

Figure 4.3. Regression Line and Scatter Plot of Number of Lane Exceedences vs. Dynamic Completion Time.

In the basic TSD model, both signal and noise are represented by an evidence variable which varies along a continuum. A diagnostic system has greater or lesser discrimination power depending on how widely separated are the probability distributions for the noise and signal distributions on this continuum. The most basic TSD model assumes that the signal and noise distributions are both gaussian with equal variances, but other distributions can be modeled, and distribution-free methods of TSD analysis are available. A highly discriminating diagnostic system or sensor separates these two distributions so there is little overlap. In a poorly discriminating diagnostic system or sensor, the signal and noise distributions overlap a great deal.

TSD also provides an independent assessments of decision bias. Whatever diagnostic sensitivity a system has, the decision maker can, indeed must, pick a decision criterion at some point along the continuum of the evidence variable. Once selected, if the evidence is beyond this criterion in one direction, the decision will be that a signal is present. Conversely, if the evidence is on the other side of the criterion, the decision will be that no signal or only noise is present. Once a decision is reached, the outcomes can be characterized in one of four categories:

- True Positives (TP): The decision is that a Signal is present and indeed this is true;
- False Positives (FP): The decision is that a Signal is present and this is, in fact, false;
- True Negatives (TN): The decision is that no Signal is present and indeed this is true;
- False Negatives (FN): The decision is that no Signal is present and this is, in fact, false.

For a system of fixed diagnostic sensitivity, true positives and false positives rise and fall together as the criterion is moved up and down the continuum of the evidence variable. For a system of fixed sensitivity, true positives cannot be increased without increasing the number of false alarms also. False alarms cannot be decreased without a corresponding decrease in the number of true positives or correct detections.

TSD provides guidance for selection of the optimal criterion based on expected value theory. If the decision criterion is termed beta and symbolized by β, then an optimal criterion can be specified as follows (Swets, 1992):

$$\beta = \frac{(V_{TN} - C_{FP})}{(V_{TP} - C_{FN})} \times \frac{P(n)}{P(s)}$$

where

β is the decision criterion expressed as a likelihood ratio in a signal detection sense, i.e., β = P(x| Signal)/P(x| Noise) where x is the evidence variable
V_{TN} is the value of making a true negative judgement or correct rejection
C_{FP} is the cost of making a false positive judgement or false alarm
$P(n)$ is the probability of noise (i.e., no signal)
V_{TP} is the value of making a true positive judgement or hit
C_{FN} is the cost of making a false negative judgement or miss
$P(s)$ is the probability of signal (i.e., true hazard)

The selection of a criterion based on expected value has much to commend it as a rational method of decision making. However, at present the author will not attempt to estimate the various values needed to exercise the preceding equation. Swets (1992) acknowledges that such estimates can be very difficult to make. He goes on to suggest that it may be enough to consider only a cost-benefit ratio, without estimating the absolute values of each term and without separating numerator and denominator in the equation above. Suffice it to say that in an attempt to capture as many signals as possible or maximize true positives, there will be an associated increase in the number of false positives. Each has its associated costs and benefits.

For a given diagnostic system, its signal detection performance can be graphically portrayed with a unit-square plot of false positive probabilities along the x-axis plotted against true positive probabilities on the y-axis. Points on this plot, termed a receiver operating characteristic (ROC) curve, are generated by varying the decision criterion or threshold several times. The most stringent criterion would plot in the lower left corner (i.e., no false positive judgements but no true positive judgements either). The most lax criterion would plot in the upper right corner (i.e., true positive probability of 1.0 with a false positive probability of 1.0 as well). A diagonal line that connects these two extremes connotes noise and signal distributions that completely overlap. ROC curves that bow above the diagonal toward the upper left connote increasingly sensitive or discriminating diagnostic systems, independent of the decision criteria used. Put another way, the greater the separation of noise and signal distributions along the continuum of the evidence variable, the greater will be the area under

the ROC curve. Given the interpretation of the diagonal and the unit-square area, the area under the ROC curve varies from 0.5 (indicating no diagnostic sensitivity or power to distinguish signal from noise) to1.0 (indicating virtually complete separation between signal distribution and noise distribution).

It was explained earlier that the ROC curve for a given diagnostic system is generated by varying the criterion along several values and plotting the resulting true positive and false positive pairs. Based on a set of N discrete criteria, the area under the ROC curve may be approximated by a trapezoidal rule (McNicol, 1972, pp. 113 - 115). Recall that a trapezoid is a four-sided figure with two sides parallel. The area of a trapezoid is equal to one-half the distance between the parallel sides times the sum of the lengths of those parallel sides. Applying this area rule to the ROC curve plotted by varying the criterion or decision N ways, the area under the ROC curve is approximately:

$$P(A) = \frac{1}{2}\sum_{i=1}^{N+1}[P(FP_i) - P(FP_{i-1})]*[P(TP_i) + P(TP_{i-1})]$$

where

$P(A)$ = the area under the ROC curve, with $0.5 \leq P(A) \leq 1.0$
$P(FP_i)$ = the false positive rate or probability associated with criterion level i
$P(FP_{i-1})$ = the false positive rate or probability associated with criterion level i -1, more lax than i
$P(TP_i)$ = the true positive rate or probability associated with criterion level i
$P(TP_{i-1})$ = the true positive rate or probability associated with criterion level i -1, more lax than i

In order to proceed with the TSD analysis, classification performance of the 15-second rule was analyzed with the following hazard criteria. The presence of "degraded driving" first had to be operationally defined somehow. It was operationally defined as 2 or more drivers exhibiting 2 or more lane departures during task execution on the 7.5 mile test track. This operational definition is somewhat arbitrary but relies on two assumptions. The first assumption is that lane departures are safety-relevant. This assumption rests on notions presented in Tijerina, Kiger, Rockwell, and Wierwille (1996) that lane departures are the first event to precipitate various types of crashes, including lane change crashes, road departure crashes, and opposite direction crashes. The second assumption behind the operational definition is that anyone might have a single lane departure during task execution that is unrelated to the task. While the absence of lane departures for some task performances on the TRC 7.5 mile test track reported in Section 2.0 of this report indicate that errorless performance is possible, the operational definition errs on the side of leniency. Thus, more than one lane departure and more than a single driver must be associated with a given task to classify that task as associated with "degraded driving."

Having defined the real-world criteria, the classification scheme applied was operationally defined that if 2 or more drivers had static completion times of greater than approximately 15 seconds, the task would fail the test. At the time the study was conducted, the actual criteria to be applied were in J2364 were in flux. The classification scheme was created again an attempt to err on the side of leniency. That is, it is possible that any given test participant out of the ten might, purely by chance, have a poor showing on a static test. On the other hand, the finding of more 2 or more out of 10 test

participants taking longer than the criterion time was interpreted to be indicative of a more enduring problem in task performance.

Consider first the area under the ROC curve as an estimate of the sensitivity of static completion time as a predictor of lane exceedences during task completion while driving. In the present case, the diagnostic system's evidence variable is static completion time. The criterion is the 15-second rule, i.e., any task which takes longer than 15 seconds to complete fails and any task which takes 15 seconds or less to complete passes. In order to generate a sensitivity measure, the criterion was varied and the classification of dynamic trial outcomes (in terms of lane exceedences as previously explained) were retabulated each time. Table 4-1 presents a range of criterion values, the resulting TP and FP values, and the calculation of P(A). Figure 4.4 depicts the ROC curve obtained from plotting the TP and FP pairs from Table 4-1, together with the diagonal that represents a complete lack of discriminative power. It can be seen from Table 4-1 that the area under the ROC curve is estimated to be 0.55. Given that the lower limit of this metric is 0.5 (corresponding to the area under the diagonal), this is taken to indicate that the static completion time has diagnostic sensitivity that is almost nil.

Table 4-1. Calculation of the Area Under the ROC Curve for the Static Completion Time Measure.

Index N=4	**Static Time Criterion**	**FP Rate**	**TP Rate**
i=4+1	10 sec	6/6 (or 1.00)	9/9 (or 1.00)
i=4	15 sec	5/6 (or 0.83)	8/9 (or 0.89)
i=3	20 sec	5/6 (or 0.83)	8/9 (or 0.89)
i=2	25 sec	4/6 (or 0.67)	7/9 (or 0.78)
i=1	30 sec	4/6 (or 0.67)	7/9 (or .078)
i=0	∞ sec	0/6 (or 0.00)	0/9 (or 0.00)

Calculation of Area Under ROC Curve, P(A):

$$P(A) = \frac{1}{2}\sum_{i=1}^{N+1}[P(FP_i) - P(FP_{i-1})]*[P(TP_i) + P(TP_{i-1})]$$

P(A) ≈ ½ [(0.67 - 0.00)(0.78 + 0.00)
+ (0.67 - 0.67)(0.78 + 0.78)
+ (0.83 - 0.67)(0.89 + 0.78)
+ (0.83 - 0.83)(0.89 + 0.89)
+ (1.00 - 0.83)(1.00 + 0.89)]
≈ ½ [1.11] ≈ 0.55

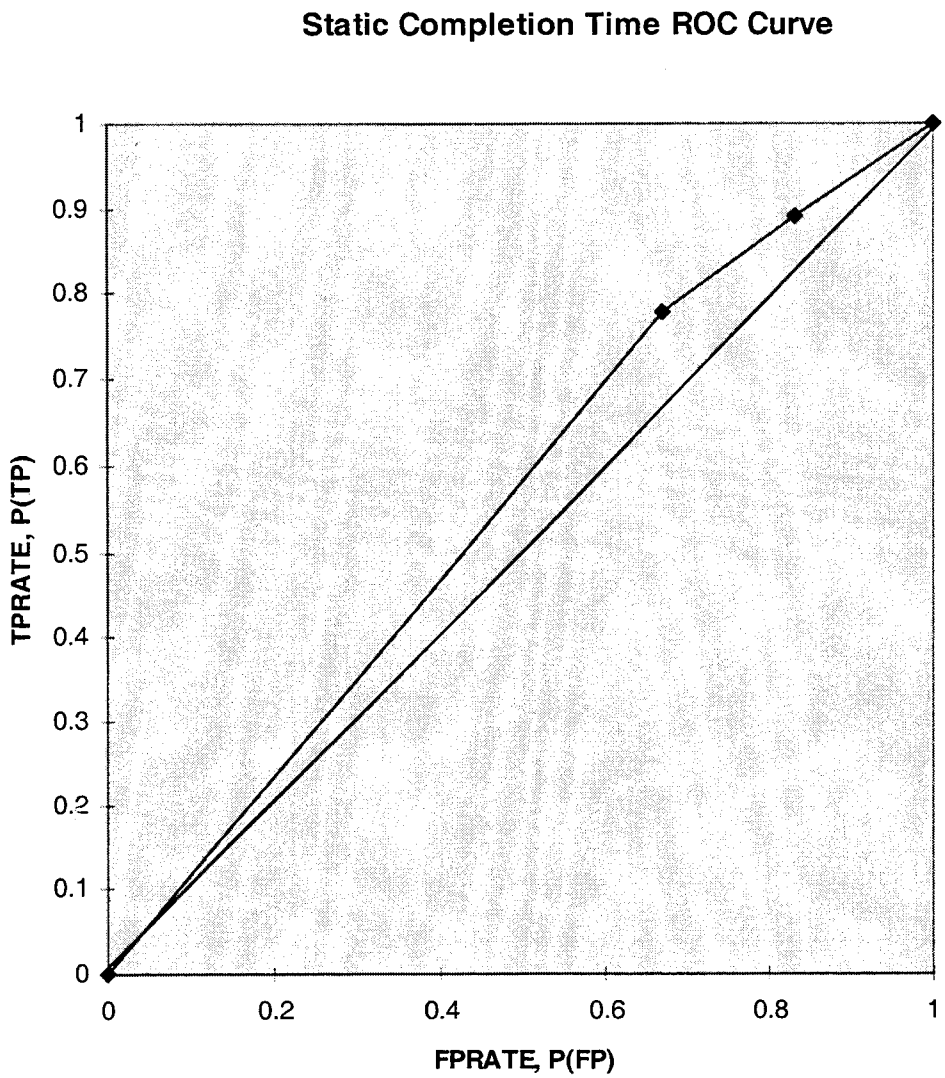


Figure 4.4. Static Completion Time ROC Curve. (Diagonal represents zero discriminative power in diagnosis or classification).

Next consider the classification results for the 15 tasks tested with the 15-second criterion. In terms of True Positives, all route navigation system destination entry tasks that required visual-manual methods both failed the 15-second rule and were associated with disrupted lanekeeping; manually dialing an unfamiliar 10-digit phone number on the cellular telephone also fell into this classification. In terms of False Negatives, tuning the Clarion after-market radio took less than about 15 seconds for static completion yet was associated with above-threshold disruptions of lanekeeping on the test track. In terms of True Negatives, the HVAC adjustment was the only task which was both completed in less than 15 seconds statically and had no appreciable effect on lanekeeping during the test track trials. Finally, of the 15 tasks, 5 of them were categorized as False Positives: dialing home on a

cellular telephone, all VAAN destination entries by voice, and tuning the after-market radio to a prescribed FM station. These five tasks all took longer than 15 seconds to complete in a parked vehicle (i.e., statically) yet were not associated with significant disruptions in lanekeeping on the test track.

Table 4-2 indicates the four-fold classification of True Negatives, False Negatives, False Positives, and True Positives from the current analysis. The classifications in Table 2 made use of the 15-second criterion with the rules of application described earlier. Of the 15 tasks examined, 9 were correctly classified and 6 were incorrectly classified. Each category of results will be discussed in turn.

Table 4-2. Classification Results of 15-Second Rule.

		Lanekeeping Problems on Test Track?	
		NO	**YES**
15 Second Rule Violated?	**NO**	*True Negatives* HVAC Adjust	*False Negatives* Radio Tune - AM
	YES	*False Positives* Cell Phone Dial - Home VAAN - Address Entry VAAN - Intersection Entry VAAN - POI Entry Radio Tune - FM	*True Positives* Cell Phone Dial - Unknown Alpine - Address Entry Alpine - Intersection Entry Alpine - POI Entry Delco - POI Entry Zexel - Address Entry Zexel - Intersection Entry Zexel - POI Entry

The correct classifications included the HVAC adjustment, all visual-manual destination entry tasks, and manually dialing a unknown 10-digit number into a cellular telephone. The HVAC task was the only true negative, meaning it was completed under 15 seconds and was associated with no significant disruption in lane keeping performance. On the other hand, the True Positives were tasks that took longer than 15 seconds to complete statically and were associated with degraded lane keeping performance on the test track. These included all destination entry tasks completed through visual-manual methods, i.e., with the Alpine, Delco, and Zexel systems, as well as the manual dialing of an unfamiliar number on the cellular telephone.

Figure 4.5 shows the scatter plot of static completion times as a function of the type of device being used. It is interesting to note that the static completion times for the destination entry tasks are generally very long. The shortest static completion time is around 40 seconds, while the majority of

times are well above 100 seconds. Keep in mind that these are completion times while sitting in a parked car concentrating on the destination task alone! From this, one can infer that beyond a 15 second rule, a 30 second rule or a 45 second rule would have classified these egregiously difficult tasks about as well.

Consider next incorrect classifications. These include both False Negative and False Positive cases. The False Negative case was tuning the Clarion aftermarket radio to an AM frequency. This took less than 15 seconds when performed statically yet was associated with disrupted lane keeping performance when done concurrently with driving. The remaining incorrect classifications were False Positives. These included all of the VAAN voice-recognition system destination entries, manually dialing home on the cellular telephone, and tuning the aftermarket radio to an FM frequency. In the case of a False Positive, the static completion time was greater than 15 seconds, yet there was no evidence of degraded lane keeping performance (as operationally defined earlier). Some salient points about these mis-classifications are discussed below.

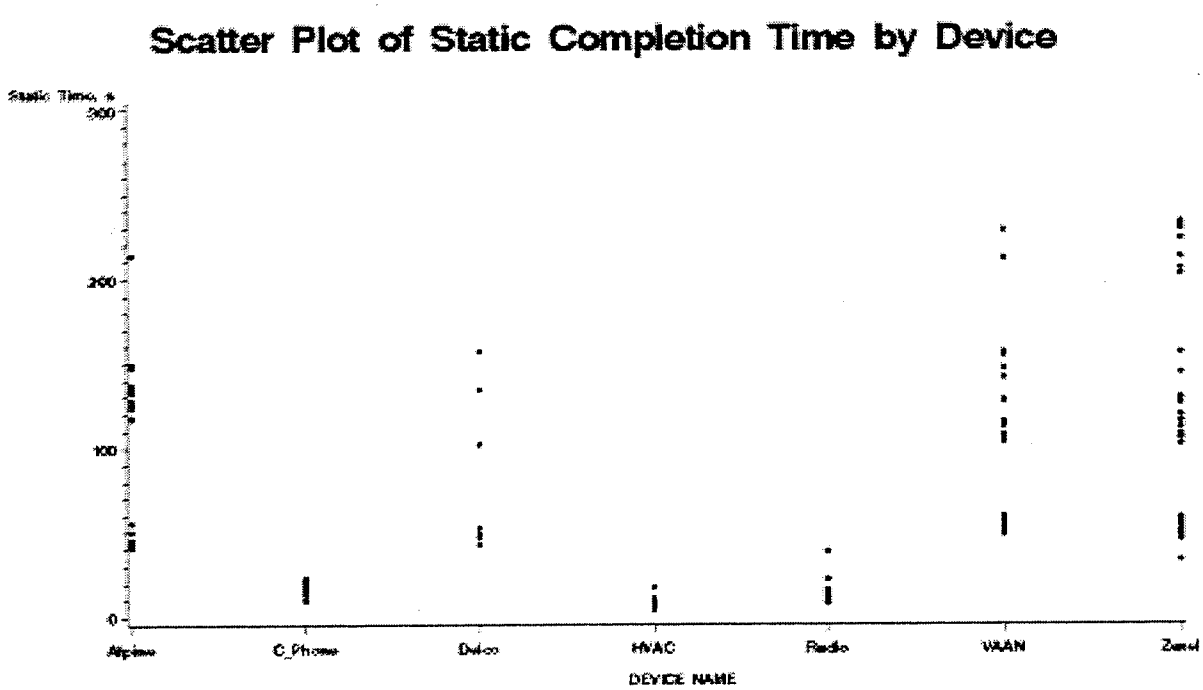


Figure 4.5. Scatter Plot of Static Completion Times For Each Device.

4.4 Discussion

The results of this test of the 15-second Rule will be discussed in terms of completion times, lane keeping performance, and classification performance. A review of what the 15-second rule does not address will then be described in broad terms. This discussion will serve as a backdrop to recommendations that conclude this report.

4.4.1 On the Nature of Completion Times: Three striking results relate to the completion times, both static and dynamic, that were observed in this study. First, the correlation between static and dynamic completion times proved to be only modest, leaving 61% of the variability in dynamic completion time unaccounted for by knowing static completion time. Second, Figure 4.1 illustrated the counterintuitive finding that a given task sometimes took less time to complete while currently driving than it took while performed singly in a parked car. Third, it was surprising that seemingly simple tasks like manually dialing a cellular telephone or tuning an aftermarket radio often took longer than 15 seconds to complete when done statically in a parked vehicle.

The reason for these results appears to lie primarily in the stochastic nature of errors that occur while completing a task. The completion times included errors and error recovery that arose during task execution. No apparatus was available to record the fine structure of such errors but that was not the point of the study. A structured walk-through by a human factors professional would identify several product design features that contribute to error occurrences. Observation of task performance would verify and amplify on the results of the structured walk-through. From the standpoint of evaluation for accessibility of a device function while driving, the error-inducing properties of a given system are properly considered part and parcel of the safety-relevant distraction potential posed. Errors are not an evaluation problem, they are a problem for evaluation.

In a related vein, the long completion times associated with even static, single-task performance is indicative of the impact these new technologies can have on drivers, especially older drivers. It is true that the test participants in this study were not long-time users of any of the devices chosen for testing. It is well known that, given extensive practice, humans can develop high levels of dual-task performance. Indeed it is such extensive practice that, for example, allows a professional musician to listen to a request while concurrently playing a piece. In the automotive safety context, however, the possibility of such levels of performance are less than satisfying. On the road, one must survive long enough to complete the "practice" sufficient to obtain dual task finesse, parallel processing, or automaticity (Shiffren, 1988). Performance failures on the primary driving task are the crux of the highway safety concern.

4.4.2 Lane Keeping Performance and Completion Times: Green (1998) carried out an ingenious meta-analysis of published studies and data relating various measures of task performance to lane keeping performance. In particular, he found a positive correlation between dynamic task completion time and number of lane exceedences or lane departures during task execution. The proportion of variability in lane exceedence counts that was accounted for by dynamic completion time varied between 48% and 56%, slightly better than the current study's 43%. However, when using static completion time to predict the number of lane exceedences, the proportion of variability accounted for dropped to only 27% in the present study. This is attributed to the large variability associated with error occurrences and error recovery during task completion.

It has been mentioned by reviewers of this study that it would have been useful to collect a baseline of lane exceedences on the TRC 7.5 mile track, i.e., of driving on the straightaways without any in-vehicle task. Data from the present study do not allow for absolute comparisons of task effects from those associated with driving alone. This was not considered an important point at the time the study was designed. The study described in Section 2.0 of this report was also conducted on the TRC 7.5 mile track used the same set of tasks and similar procedures of training with a different group of 16 test participants. Half of the test participants in that study were younger (ages 18 to 25 years) and half were older (ages 55 to 65 years). Those test participants demonstrated that it was possible to complete the VAAN tasks without any lane exceedences. This suggests that an expected value of zero lane exceedences on a 2.5 mile straightaway on the TRC track when traveling at 45 mph is not unreasonable and degraded lane keeping might reasonably be attributed to in-vehicle task execution. This in turn implies that most if not all of the lane exceedences observed may rightly be attributed to the concurrent in-vehicle task rather than random variation in driving. It is acknowledged that such an interpretation may not be true in open road driving where lane exceedences of various types are more or less common (e.g., curve cutting).

4.4.3 Classification Accuracy of the 15-Second Rule: The TSD analysis illustrated that the diagnostic sensitivity of the static completion time measure is close to nil. That is, the noise (not distracting) and signal (distracting) distributions overlap a great deal on this evidence variable. As a result, the selection of 15 seconds for static completion time has associated with it both false negatives and false positives. False negatives are of safety concern because, if SAE J2364 were implemented broadly, some functions would be deemed suitable for accessibility while driving that have a demonstrated capacity to disrupt the driver's ability to maintain the vehicle in the travel lane. In the case of the tuning of the after-market radio, it was placed in the dashboard, low and toward the centerline of the subject vehicle. Test participants therefore had to reach over to acquire the small "band" and "seek" buttons needed to accomplish the task. When done in a parked car, the test participants were able to accomplish the AM tuning task within the 15 second criterion. When driving, lane keeping was disrupted, perhaps because of bias on the steering wheel as people reached over to press the necessary buttons. Thus, false negatives should be of concern for safety reasons.

The false positives should be of concern for reasons other than safety. In these instances, test participants were not able to complete a given task within the 15 second deadline yet there was no significant degradation in driving performance, as operationally defined in the study. This implies that, if such tasks (or systems) were being evaluated solely on the basis of J2364, they would fail needlessly. Thus, designers would waste time attempting to correct a non-problem. The economic costs of such false positives are unknown but are presumed to be non-zero in the competitive and time-sensitive field of automotive products.

4.5 Recommendations

The primary motivation behind the 15-Second Rule was to support highway safety. The goal was to develop an inexpensive and easy-to-use method to assess whether or not a route navigation system function may be accessible to the driver while the vehicle is in motion. The selection of static completion time as the measurement system evolved from a literature review and re-analysis of published data sets, a chain of causal inference, and expert judgement of the SAE Safety and Human

Factors sub-committee members working on SAE J2364. The literature review and re-analysis of published data sources is contained in Green (1998) and was thoughtfully carried out. The chain of causal inference deduced from Green's analysis to support SAE J2364 went something like this. Lane exceedences are considered a *prima facie* safety-relevant driving performance measure. Lane exceedences are positively correlated with and logically relatable to the number of eye glances to the in-vehicle device. The number of eye glances to the in-vehicle device during task execution are positively correlated with overall dynamic completion time. Dynamic completion time is positively correlated with static completion time. *Ergo*, static completion time must be positively correlated with lane exceedences, i.e., with safety.

To this chain of causal inference were added the following considerations. Empirical measurement of driver eye glance behavior is difficult and time-consuming to accomplish. For this reason, the SAE S&HF sub-committee assumed eye glance measurement would not be practical for device evaluation. Capturing driving performance while completing an in-vehicle task was also assumed by the subcommittee to be too onerous for practicable testing. It was presumed that such testing requires an instrumented vehicle, associated signal processing and data reduction capabilities, access to a test track, and trained personnel. Such resources can be hard to come by. Static completion time was compatible with GOMS modeling as advocated in the companion standard, SAE J2365. GOMS modeling techniques have so far only been developed for single task performance, not dual-task time sharing performance. This necessitates a criterion of completion times for tasks performed singly, i.e., statically. Finally, the proliferation of devices of untrammeled complexity finding their way into cars and trucks created a sense of urgency. The sub-committee perceived a need to provide a practicable method sooner rather than later, to help stem the tide of potentially dangerous driver distractions. In this regard, the 15-second rule was considered substantially better than no standard at all.

A set of recommendations may be generated by analyzing the assumptions and line of causal inference just described. This is done in the sections below.

4.5.1 Necessary and Sufficient Safety-Relevant Measures of Driving Performance: It is commonly assumed that lane exceedences during task execution are fundamentally safety relevant (e.g., Tijerina, Kiger, Rockwell, and Wierwille, 1996). To many, it seems self-evident that the driver must control the vehicle and remain in the travel lane, moving from it only in a controlled fashion. Failure to properly keep in one's lane is the proximal event that leads to such crash types as single vehicle road departures, lane change crashes, and opposite direction crashes. Thus, it appears that lane departures are directly safety-relevant. Two interesting questions arise, however. First, is a lane exceedence really fundamentally safety-relevant? Second, is a driver-vehicle performance measure like lanekeeping sufficient to establish a safety link?

It is useful to question whether lane exceedences or lane departures are sufficient to capture the notion of "degraded driving" and, if not, what might be done to usefully operationalize the concept further. Brown (1994) has pointed out that driver distraction may take at least two distinct forms which he refers to as a general withdrawal of attention and a selective withdrawal of attention. The general withdrawal of attention manifests itself with degraded vehicle control and degraded object and event detection. This category of distraction is primarily linked with physical processes such as eyelid closure (in the case of fatigue) and eye glances away from the driving scene. The selective

withdrawal of attention is more insidious in that it is associated with degraded object and event detection while leaving vehicle control largely unaffected. The mechanisms underlying the selective withdrawal of attention are those associated with cognitive load more than physical processes. For example, selective withdrawal of attention might arise event though the driver is looking at the road ahead due to attention to inner thoughts or other tasks as indicated by open-loop (expectancy-driven) as opposed to closed-loop (stimulus-driven) visual scanning, restricted sampling of mirrors and the road scene, and a type of empty field myopia. Thus, it appears that while measures like lane keeping, speed maintenance, car following performance, and the like are appropriate, they are not sufficient for a fuller evaluation of driver distraction. At a minimum, it is recommended that future research and methodology also incorporate object and event detection performance such as driver reactions to lead vehicle braking, incursion into the travel lane, or response to a sudden road curvature, road hazard, or other event.

4.5.2 Beware The Chain of Causal Inference: Care must be taken in attempts to build predictions from chains in correlations. Consider that correlations exist between measures A and B and correlations exist between measures B and C. Intuitively, it seems reasonable to use such correlational information to predict the correlation between A and C. However, Tijerina et al. (1996) demonstrated by means of Venn diagrams that even if the correlations between A and B and B and C are as high as 0.7 (and this is unusually high for human factors studies), the correlation between A and C turns out to be indeterminate, with a range anywhere between perfect (1.0) to no correlation (0.0). Therefore, it is recommended that additional measures beyond static completion time be taken. Examples of additional measures that might be taken during the accessibility evaluation process might include dynamic completion time; vehicle control measures like lane exceedences, speed variance, or time headway variability; object and event detection measures like brake reaction time or minimum Time-to-Contact; and subjective assessments by drivers; and subject matter expert (SME) checklist evaluations of the interface. Additional empirical measures which support the same conclusion provide what Garner, Hake, and Eriksen (1956) termed "converging operations" and this adds confidence that a meaningful, comprehensive, and coherent assessment can be achieved.

4.5.3 Eye Glance Measurement is Not Necessarily Too Difficult: Eye glance behavior measurement need not be too difficult to collect. Green (1998) found that, as compared to the number of glances to a device, eye glance duration to a device was a poor predictor of lane keeping performance. Indeed, there are reasons why the range of eye glance durations is generally kept narrow while driving (Rockwell, 1987; Wierwille, 1993). It would be a simpler process to only count the number of glances to a device rather than to measure the duration of each glance, the state transitions from one glance location (e.g., road ahead) to another (e.g., device), or the percentage of time spent gazing at a given location (e.g., road ahead). Furthermore, SAE J2396 (Society of Automotive Engineers, 1999) has been drafted to provide a recommended practice for eye glance behavior measurement that requires only a video camera and a recording device. Thus, it is recommended that an analysis of the effort required to empirically collect only the number of glances to a device be considered for inclusion in a revised measure. Furthermore, it is recommended that additional correlational analysis be carried out to verify the robustness of the models in Green (1998) with new data sets.

4.5.4 Efficient Driver Performance Measurement Is Feasible: The assumption that capturing driver performance is too difficult is also questionable. The present study, for example, did not use

an instrumented vehicle. A ride-along experimenter with a stop watch and an observer in a following vehicle were all that were required to provide the data analyzed here. It is recommended, therefore, that a driving test be incorporated into device evaluation and revisions of J2364.

4.5.5 Static Completion Time May be Highly Misleading as a Link to GOMS Modeling: Errors are the key in this regard. It is errors that are the root cause of the high variability found in static completion times and dynamic completion times. Recall that GOMS models assume error-free performance accomplished by means of a known strategy. The results of the present study indicate that errors are quite common over a representative range of in-vehicle tasks with commercially available devices. Once an error occurs, error recovery does not follow a single path to solution, especially among the "advanced novices" that are emulated in the testing reported here. These facts pose the possibility that the correlation between static completion time and GOMS estimates will be low. Beyond this, errors brought on by poor device design are a proper focus of evaluation in their own right. It is inappropriate to consider errors as a pesky measurement to be avoided, removed by simplifying assumption, or otherwise ignored. The etiology of errors and error recovery should be characterized during safety evaluations and rectified in iterative design. It is recommended, then, that GOMS modeling may be used as a computational design tool early in the product design process. However, GOMS modeling must be augmented as early as feasible by empirical data collection in the product development cycle.

4.5.6 An Imperfect Evaluation Rule is Better than Nothing Because Poor Driver Interfaces are Threatening Highway Safety: This argument has been put forth in support of the 15-Second Rule by certain members of the SAE Safety and Human Factors Subcommittee. Indeed, it was the existence of commercially available route navigation systems with access to very demanding functions that motivated the project from the start. The fact that the study reported here was conducted on a test track was due to concerns by the project staff that it was simply too dangerous to perform on the open highway. A most sobering experience is to have a driver attempt such complex tasks while driving on the highway... when you are sitting in the passenger seat.

Reaction to this perceived urgency is tempered somewhat by the lack of empirical data to demonstrate a relationship between destination entry and crash involvement. Certainly, it is plausible that poor interfaces that place too great a demand on the driver may precipitate crashes but where are the numbers to demonstrate the point? As far as we know, there is no epidemic of crashes that can be attributed to route guidance systems, though this may be due to shortcomings in crash reporting (cf., Goodman et al., 1999). Supporters of SAE J2364 would contend that the safety and human factors community must be proactive to prevent such a crash epidemic from arising in the first place. Evidence of distraction-related crashes is accumulating both here and abroad that bolster the argument for some stopgap evaluation method to promote highway safety. For example, the Japanese National Policy Agency (1998) is tabulating crash statistics that implicate route navigation system interactions as crash contributors in at least some instances.

On the other side of the debate are those who argue that the science behind the 15-Second Rule is specious. Many members of the SAE Safety and Human Factors Committee are concerned over the seemingly poor discriminative power of the 15-Second Rule. Adoption of SAE J2364 would, in their view, introduce a scientifically questionable rule that designers will strive to meet, with potentially disastrous results. To this criticism, defenders of SAE J2364 caution against making "... the perfect

the enemy of the good." In a related vein, Norman (1998) recently cautioned the human factors community to learn when "... good enough is good enough", or else risk becoming irrelevant. Finally, Abelson (1995) recently provided a humorous yet cogent reminder of where the behavioral sciences (and, by inference, human factors) stand relative to physicists on the continuum of measurement accuracy. As he put it, the anxieties physicists have over randomness is exemplified by worrying about atomic clocks drifting by 0.01 seconds per century. On the other hand, behavioral scientists typically operate at a level of measurement precision more akin to an alarm clock that gains or loses 6 hours per week.

Consider but two of the more salient points of contention. One has to do with the driver's ability to "chunk" the in-vehicle task at will. Multiple glances of various glance durations between the device and the roadway are critical to in-vehicle task completion while concurrently driving. Indeed, the number of glances and glance durations are a function of a number of factors including driving conditions (e.g., traffic density, speed, weather), system location, individual differences, and systems design. Thus the "chunking" or partitioning of a task (number of glances) represents the willingness of a particular user to divert visual attention from the roadway under a given set of circumstances. Since the only opportunity to influence chunking within the standard is in system design, it is vital that any evaluation procedure address this issue. Indeed, two systems with identical task completion times may exhibit very different time-based demands and consequently very different task chunking. This constitutes a serious unknown of the proposed rule.

Consider a second point of contention. It may be possible to meet the 15-Second Rule with a design that aggravates crash risk rather than alleviates it. Is it possible to design an interface function to meet a 15-second static completion time rule that actually makes dynamic performance worse?

An experiment to test this hypothesis might be constructed in the following manner. First, one or more tasks would be developed that minimize static completion times to below 15-seconds. However, this set of tasks would might have properties that, through considerations beyond the 15-Second Rule, might lead to problems during dynamic execution or timesharing while driving. Second, a set of tasks would also be prepared that did not necessarily minimize completion time or even meet the 15-Second Rule but, again through other considerations, were thought to have properties that might make such a transaction acceptable to perform while driving. The critical analysis would involve, over an ensemble of test participants, assessment of the false positive and false negative rates of both categories of design. An outcome supportive of the concern would be one where the contrary dynamic performance results were confirmed in test track or on-road testing.

Hierarchical menu design appears to be one area both relevant to modern automotive device design and ripe for application to the investigation. Menus represent one of the most popular forms of human-computer interface, yet relatively little is known about their application in automotive settings (i.e., while timesharing item search with driving) as opposed to office settings (i.e., while concentrating on item search on a larger screen, without a concurrent continuous control task, etc.). Miller (1980; 1981) conducted an early investigation of the menu breadth-versus-depth tradeoff. His database consisted of a hierarchical tree structure of 64 items that depicted the superordinate/subordinate relations among items. The database was constructed such that it could be represented with any of four different menu structures:

(a) menus with breadth of two choices at each of 6 levels of depth (2^6);
(b) menus with breadth of four choices at each of 3 levels of depth (3^4);
(c) menus with breadth of eight choices at each of 2 levels of depth (8^2); and
(d) a menu with 64 choices presented on a single level of depth , i.e., on a single display (64^1).

Target items were always at the lowest level of the hierarchical tree. Miller (1981) measured total time to find each target item and also recorded the probability of not reaching the target before a timeout of 10 seconds occurred. The system response time to move from one menu screen to the next was fixed at 0.5 seconds. His results indicated that test participants were slower and less accurate in finding the target items with structures (a) and (d). Test participants were faster and more accurate with the 8^2 structure. Miller (1980) concluded that for systems of moderate size, system breadth is preferable to system depth. These results were confirmed using a different 64-item database by Kiger (1984), who argued that the design option of presenting 8 options on a each of two levels of a menu hierarchy was compatible with the characteristics of human short term memory. However, Snowberry, Parkinson, and Sisson (1983) were able to demonstrate the beneficial trend of greater menu breadth over less depth even to a single screen of up to 64 items. This could be achieved if, unlike the single 64-item displays of Miller (1980) and Kiger (1984), a strict categorical grouping was maintained in the single menu. Subsequent research has supported the general design guidance that, for typical office settings, breadth is preferred over depth, at least until the point of diminishing returns caused by a cluttered display.

Now, assume that a designer attempts to follow this hierarchical menu design guidance for an automotive application. Assuming a static task (i.e., no concurrent task to timeshare with), the expectation is that a lot of information presented on a single visual display will produce shorter target search times when compared to putting a smaller breadth of options over two or more levels of a hierarchical menu. Human vision is quite powerful and is one reason why physical item matching is faster than trying to match categories. Furthermore, transition from menu to menu will involve some system response time that will in all likelihood be more time consuming than saccadic eye movements across a single screen of options. Therefore, if a system designer was familiar with the menu breadth-over-depth design guideline, he or she might apply it in good faith and gain a sense of "validation" from from a static evaluation via the SAE 15-Second Rule.

The important question is what will happen while driving. Is there any reason to believe the benefits of breadth over depth of hierarchical menu design would not apply while driving? No studies of the type described here have been conducted yet. However, there is reason to believe that the design guideline might not work well while concurrently driving. For example, a dense screen will be harder for the time-sharing driver to search for and fixate on targets. This difficulty is predicted in part due to the large eye movement shifts needed to shift one's gaze from the display to the road scene and back again. Reading research has indicated that such eye movements are one reason why some readers have more difficulty than others.

Given the increasing attention by the public and media to in-vehicle technology (as evidenced by frequent public outcry associated with cellular telephone use by drivers) and the potential risks, it would *not* appear that a 15 seconds task time limit would be viewed as sensible given that it implies that a driver can safely look away from the driving scene for 15 seconds (although it is highly unlikely anyone would exhibit such behavior it is a technical possibility within the bounds of the draft

standard). This perception has the potential to diminish the credibility of SAE standards in general and human factors in particular. Without addressing task chunking (e.g., sub-task glance times and frequency of glances per task) in the standard it simply would not appear credible.

To the extent that the standard is inadequate in addressing concerns, it is likely that any "availability of function" standard implemented by SAE will evolve as a universal and influence the design of other information oriented in-vehicle technologies. Indeed, with plans to provide OEM hardware to allow access to the Internet from a moving vehicle, this would potentially give license, as currently proposed, to receive extended e-mail messages where the opportunity for "chunking" is minimized by the very nature of the service. This heightens the importance of drafting a supportable and defensible standard. If the required data does not exist, every effort should be made to develop an agreed upon process for obtaining it including a research plan acceptable to the committee

This section concludes the following summary comments and recommendations to improve on the 15-second rule.

- The limitations of the 15-second rule and the ideas behind it may yield suggested areas for improvement in the development of objective test procedures for ITS information systems.

- None of the empirical support for the 15-second rule examined speed maintenance or object and event detection performance. An example of an object or event to be detected is a lead vehicle that suddenly decelerates, an object suddenly thrust (or appearing) into the host vehicle travel lane, or a sudden change in the roadway (e.g., lane drop, construction zone) that requires prompt driver action. We propose that correlational studies be done to relate driver reaction latencies and efficiencies in response to objects and events to static and dynamic completion time. Since crash involvement is related to such lapses or delays in object and event detection, this will be a useful further evolution of the simple completion time rule.

- The 15-second rule does not address the issue of how a task might be chunked (or not chunked). We propose a checklist be developed to examine properties of an interface (e.g., character size or contrast) or system logic that thwart the voluntary chunking of a task and the reacquisition of a task from the point where the driver left off. The validation of the items in the checklist can be conducted by applying those items to several products, examining driver-vehicle behavior and performance while driving and concurrently engaging in such tasks, and assessing the item validity and reliability identify problems.

- There are no baselines for certain safety-relevant measures of merit. For example, number of lane exceedences that might arise with no in-vehicle task would be helpful. Among the many issues to be addressed in this context is the fact that lane exceedences while concurrently working with an in-vehicle device on a test track may be termed 'unintended', there are many driver strategies that intentionally produce lane exceedences on the open road (e.g, curve cutting is but one example). Nonetheless, there is a need to determine a threshold beyond which there is reason to be concerned.

4.6 Postscript

Since this study was conducted, SAE J2364 was modified somewhat and brought out for committee ballot to the SAE Safety and Human Factors Committee. It passed by the slimmest of margins...one vote. Numerous comments were provided and it was agreed by the chairman of the committee that these comments would be responded to, that SAE J2364 would be revised to reflect selected inputs obtained during the balloting process, and that the revised version would be resubmitted for ballot. It was subsequently decided by the committee to submit the initial ballot results to the SAE Division without further modification to the recommended practice. Most recently, the approval was overturned by the SAE Division and the proposal was returned to the Safety and Human Factors Committee for reconsideration.

5.0 References

Abelson, R. P. (1995). *Statistics and principled argument* (p. 21). Hillsdale, NJ: Lawrence Erlbaum.

Abelson, R. P. (1995). *Statistics and principled argument*. Mahwah, NJ: Erlbaum.

Akamatsu, M., Yoshioka, M., Imacho, N., Diamon, T., and Kawashima, H. (1997). Analysis of driving a car with a navigation system in an urban area. In Y. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces* (pp. 85-96). Mawah, NJ: Lawrence Erlbaum.

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis and Prevention*, Vol. 27 (5), pp. 707-715.

Alm, H., & Nilsson, L. (1994). Changes in driver behaviour as a function of handsfree mobile telephones. *Accident Analysis and Prevention*, Vol. 26, pp. 441-451.

Alm, H., & Nilsson, L. (1990). *Changes in driver behaviour as a function of handsfree mobile telephones: a simulator study*. Report No. 47, DRIVE Project V1017 (BERTIE), October. Also paper No. 175, 1991 (ISSN 0347-6049). Linkoping, Sweden: Swedish Road and Traffic Research Institute.

Antin, J. A., Dingus, T. A., Hulse, M. C., & Wierwille, W. W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display. *International Journal of Man-Machine Studies*, *33*, 581-594.

Bardy, B., & Laurent, M. (1991). Visual cues and attention demand in locomotor positioning. Perceptual and Motor Skills, 72. 915-926.

Barickman, F. (1998). Intelligent data acquisition for intelligent transportation research (SAE Technical Paper No. 981198). Warrendale, PA: Society of Automotive Engineers

Boase, M., Hannigan, S., & Porter, J. M. (1988). Sorry, can't talk now... just overtaking a lorry: the definition and experimentation investigation of the problem of driving and hands-free car phone use. In E. D. Megaw (Ed.). *Contemporary Ergonomics* (pp. 527-523). London: Taylor and Francis.

Briem, V., & Hedman, L. R. (1995). Behavioral effects of mobile telephone use during simulated driving. *Ergonomics,* Vol. 38 (12), pp. 2536-2562.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention,* Vol. 23 (4), pp. 309-316.

Brouwer, W., Waterink, W., Van Wolffelaar, P., & Rothengatter, T. (1991) Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. Human Factors, 33 (5). 573-582.

Brown, I. D., Tickner, A. H., & Simmonds, D. C. V. (1969). Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology,* Vol. 53 (5), pp. 419-424.

Burnett, G. E., & Joyner, S. M. (1993). An investigation on the man machine interfaces to existing route guidance systems. *Proceedings of the IEEE-IEE vehicle navigation and information systems conference*, 395-400. Piscataway, NJ: Institute of Electrical and Electronic Engineers.

Brown, I. D., (1994). Driver fatigue. *Human Factors, 36*(2), 298-314.

Campbell, J. L., Carney, C., & Kantowitz, B. H. (1997). *Advanced traveler information system (ATIS) and commercial vehicle operations (CVO) components of the Intelligent Vehicle Highway System (IVHS): Draft Human Factors Design Guidelines for ATIS/CVO* (Contract No. DTFH61-92-C00102). Seattle, WA: Battelle.

Campbell, K. L., Joksch, H. C., and Green, P. (1996). A bridging analysis for estimating the benefits of active safety technologies (Report No. UMTRI-96-18). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Card, S. K., Moran, T. P., and Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Dingus, T. A. (1997). Empirical data needs in support of model building and benefits estimation. In ITS America (Ed.), *Modeling ITS collision avoidance system benefits: expert panel proceedings.* Washington, DC: Intelligent Transportation Society of America.

Dingus, T. A., Antin, J. F., Hulse, M. C., & Wierwille, W. W. (1989). Attentional demand requirements of an automobile moving-map navigation system. *Transportation Research, A23*(4), 301-315.

Dingus, T. A., & Hulse, M. (1993). Some human factors issues and recommendations for automobile navigation information systems. *Transportation Research, 1C*(2), 119 - 131.

Dingus, T. A. Hulse, M., Jahns, S., Alves-Foss, J., Confer, S., Rice, A., Roberts, I., Hanowski, R., & Sorenson, D. (1996, November). *Development of human factors guidelines for advanced traveler information systems and commercial vehicle operations: Literature review* (Report No. FHWA-RD-95-153). McLean, VA: Federal Highway Administration Turner-Fairbank Highway Research Center.

Dingus, T., McGehee, D., Hulse, M., Jahns, S., Manakkal, R., Mollenhauer, M., & Fleischman, R. (1995, June). *TravTek evaluation Task C_3 Camera car study* (Report No. FHWA-RD-94-076). McLean, VA: Federal Highway Administration Turner-Fairbank Highway Research Center.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. *Human Factors,* Vol. 27(2), pp. 201-207.

Elander, J., West, R., and French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. *Psychological Bulletin, 113*(2), 279-294.

Evans, L. (1991). *Traffic Safety and the Driver* (pp. 282 - 309). New York: Van Nostrand Reinhold.

Evans, L., and Wasielekski, P. (1982). Do accident involved drivers exhibit riskier everyday driving behavior? *Accident Analysis and Prevention, 14*, 57-64.

Fairclough, S. H., Ashby, M. C., Ross, T., & Parkes, A. M. (1991). Effects of handsfree telephone use on driving behaviour. *Proceedings of the ISATA Conference*, Florence, Italy, ISBN 0 947719458.

Fairclough, S. H., Ashby, M. C., & Parkes, A. M. (1993). In-vehicle displays, visual workload and usability evaluation. In A. G. Gale et al. (Eds.), *Vision in Vehicles IV* (pp. 245-254). London: Taylor and Francis.

Foley, J. P., & Hudak, M. J. (1996). Autonomous route guidance system field test. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, 887- 890. Santa Monica, CA: Human Factors and Ergonomics Society.

Fournier, B., & Stager, P. (1976) Concurrent validation of a dual-task selection test. Journal of Applied Psychology, 61 (5). 589-595.

Garber, N. J., and Gadiraju, R. (1989). Factors affecting speed variance and its influence on accidents. *Transportation Research Record, 1213*, 64-71.

Garner, W. D., Hake, H. W., and Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review, 63*, 149-159.

Glauz, W. D., Bauer, K. M., and Miglet, D. J. (1985). Expected traffic conflict rates and their use in predicting accidents. *Transportation Research Record, 1026*, 1-12.

Goodman, M., Bents, F. D., Tijerina, L., Wierwille, W. W., Lerner, N., and Benel, D. (November, 1997). *An investigation of the safety implications of wireless communications in vehicles* (Report No. DOT HS 808-635). Washington, DC: National Highway Traffic Safety Administration.

Goodman, M.J.., Tijerina, L., and Bents, F. (1999). Wireless communications from vehicles: Safe or unsafe? *Transportation Human Factors Journal, 1* (1), 3-42

Graf, P., and Torrey, J. W. (1966). Perception of phrase structure in written language. *American Psychological Association Convention Proceedings*, 83-88. Reported in J. R. Anderson, *Cognitive psychology and its implications* (Third edition) (pp. 366-367). New York: W. H. Freeman.

Green, P. (1995). Automotive techniques. In J. Weimer (Ed.), *Research Techniques in Human Engineering* (pp. 165 - 208). Englewood Cliffs, NJ: Prentice-Hall.

Green, P. (1997). *Potential safety impacts of automotive navigation systems*. Paper presented at the Automotive Land Navigation Conference, June 18, 1997.

Green, P. (1998, April). *Visual and task demands of driver information systems*. Internal working draft. Ann Arbor: University of Michigan Transportation Research Institute.

Green, P. (1999). *SAE J2364 – Navigation and route guidance function accessibility while driving* (Draft). Warrendale, PA: Society of automotive engineers.

Green, P., & George, K. (1995). When should auditory guidance systems tell drivers to turn? *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, 1072 - 1076. Santa Monica, CA: Human Factors and Ergonomics Society.

Green, P., Hoekstra, E., & Williams, M. (1993, November). *Further on-the-road tests of driver interfaces: examination of a route guidance system and a car phone* (Report No. UMTRI-93-35). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Green, P., Levison, W., Paelke, G., & Serafin, C. (1995, December). *Preliminary human factors design guidelines for driver information systems* (Report No. FHWA-RD-94-087). McLean, VA: Federal Highway Administration Turner-Fairbank Highway Research Center.

Green, P. (1997). *Potential safety impacts of automotive navigation systems.* Paper presented at the Automotive Land Navigation Conference, June 18, 1997

Green, P. (1998, April). *Visual and task demands of driver information systems.* Working draft. Ann Arbor, MI: University of Michigan Transportation Research Institute.

Green, P. (1999a). The 15-Second Rule for Driver Information Systems, *Proceedings of the ITS America Ninth Annual Meeting* , Washington, D.C.: Intelligent Transportation Society of America, CD-ROM.

Green, P. (1999b). Estimating compliance with The 15-second rule for driver-interface usability and safety. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*. CD-ROM.

Hanowski, R., Kantowitz, B., & Tijerina, L. (1995, September). *Final report — workload assessment of in-cab text message system and cellular phone use by heavy vehicle drivers in a part-task driving simulator* (Contract No. DTNH22-91-07003). Columbus, OH: Battelle Memorial Institute.

Hayes, B.C., Kurokawa, K., & Wierwille, W.W. (1989). Age-related decrements in automobile instrument panel task performance. *Proceedings of the Human Factors Society 33rd Annual Meeting*, pp. 159-163.

Hulse, M. C., Dingus, T. A., Fischer, T., and Wierwille, W. W. (1989). The influence of roadway parameters on driver perception of attentional demand. In A. Mital (Ed.), *Advances in industrial ergonomics and safety* I (pp. 451-456). London: Taylor and Francis.

Jackson, P. (1996a). Group differences in ability to use verbal route guidance and navigation instructions. In S. A. Robertson (Ed.), *Contemporary ergonomics 1996.* London: Taylor and Francis.

Jackson, P. G. (1996b). Driven to distraction: In search of better ways to present route guidance information. *Proceedings of Parliamentary Advisory Committee for Transport Safety (PACTS)*

(1996) Conference: Transport Telematics and Road Safety: How many stand to benefit? London, October 9, 1996

Jackson, P. G. (1996c). How will route guidance information affect cognitive maps? *Journal of Navigation, 49*(2), 178-186.

Japanese National Police Agency (1998). Study of injury-producing crashes during first six months of 1998. Translation provided at SAE Safety and Human Factors Committee Meeting by Mr. Tomohiro Fukomura.

Kahneman, D., Ben-Ishai, R., & Lotan, M. (1973). Relation of a test of attention to road accidents. Journal of Applied Psychology, 58 (1). 113-115.

Kames, A. J. (1978). A study of the effects of mobile telephone use and control unit design on driving performance. *IEEE Transactions on Vehicular Technology*, VT-27 (4), pp. 282-287.

Kantowitz, B. H., Hanowski, R. H., & Tijerina, L. (1996). Simulator evaluation of heavy-vehicle driver workload II: complex secondary tasks. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting,* pp. 877 - 881.

Kennedy, R. S., Silver, N. C., and Ritter, A. D. (1995). A visual test battery: Tale of two computers. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, 907 - 911.

Kennedy, R. S., Lane, N. E., Turnage, J. J., and Harm, D. L. (1997). Individual differences in visual, motor, and cognitive performances: Correlations with a simulated shuttle landing task. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, 589- 593.

Kiger, J. I. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. *The International Journal of Man-Machine Systems, 20*, 201-213.
Kintsch, W., and Van Dijk, T. A. (1978). Toward a model of text comprehension and reproduction. *Psychological Review, 86*, 363-394.

Kimura, K., Marunaka, K., Sugiura, S. (1997). Human factors considerations for automotive navigation systems -- legibility, comprehension, and voice guidance. In. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces (pp. 153-167). Mahwah, NJ: Lawrence Erlbaum.*

Llaneras, R. E., Swezey, R. W., & Brock, J. F. (1993). Human abilities and age-related changes in driving performance. *Journal of the Washington Academy of Sciences*, Vol. 83 (1), pp. 32-78.

McKnight, A. J., & McKnight, A. S. (1991, January). *The effect of cellular phone use upon driver attention.* Washington, DC: National Public Services Research Institute. (Funded by AAA Foundation for Traffic Safety.)

McKnight, A. J., & McKnight, A. S. (1992, April). *The effect of in-vehicle navigation information systems upon driver attention.* Landover, MD: National Public Services Research Institute.

McKnight, A. J., & McKnight, A. S.(1993). The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention*, Vol. 25 (3), pp. 259-265.

McLeod, P. (1977). A dual task response modality effect: support for a multiprocessor models of attention. Quarterly Journal of Experimental Psychology, 29. 651-667.

McNicol, D. (1972). *A primer of signal detection theory*. London: Allen and Unwin.

Means, L. Carpenter, J. T., Szczubleewski, F. E., Fleischman, R. N., Dingus, T. A., & Krage, M. K. (1993). Design of the TravTek auditory interface. *Transportation Research Record, 1403*, 1-6.

Miller, D. P. (1980). *Factors affecting item acquisition performance in hierarchical systems: Depth vs. breadth*. Unpublished Ph.D. dissertation. Columbus, OH: The Ohio State University.

Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. *Proceedings of the Human Factors Society 25th Annual Meeting*, 296-300.

Montgomery, D., and Peck, E . (1992). *Introduction to linear regression analysis (2nd edition)*. New York: John Wiley.

Monty, R. W. (1984). *Eye movements and driver performance with electronic automotive displays*. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Mourant, R. R., Herman, M., & Moussa-Hamouda, E. (1980). Driver looks and control location in automobiles. *Human Factors*, *22*(4), 417-425.

National Police Agency of Japan (1998, August). *Car-phone-related traffic accidents during the first half of 1998*. Handout presented at ITS America Safety and Human Factors Committee Meeting, Dublin, OH, July 28, 1999.

Nilsson, L. (1993). Behavioural research in an advanced driving simulator: Experiences of the VTI system. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting,* pp. 612-616.

Nilsson, L., & Alm, H. (1991). Elderly people and mobile telephone use – effects on driver behaviour. *Proceedings of the Conference Strategic Highway Research Program and Traffic Safety on Two Continents*. Gothenburg, Sweden, and DRIVE Project V1017 (BERTIE, Report No. 53), March 1991.

Nilsson, L., & Alm, H. (1991). *Effects of mobile telephone use on elderly drivers' behavior - including comparisons to younger drivers' behavior*. Gothenburg, Sweden, and DRIVE Project V1017 (BERTIE, Report No. 176), 1991.

Norman, D. L. (1998). Plenary address given at the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago, IL.

Older, S.J., and Spicer, B.R. (1976). Traffic conflicts - A development in accident research. *Human Factors, 18*(4), 335-350.

Pachiaudi, G., & Chapon, A. (1994). Car phone and road safety. *XIVth International Technical Conference on the Enhanced Safety of Vehicles*, No. 94-S2-0-09. Munich, Germany.

Paelke, G. (1993). A comparison of route guidance destination entry methods. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 569- 573. Santa Monica, CA: Human Factors and Ergonomics Society.

Parkes, A. M. (1991). Drivers business decision making ability whilst using carphones. In Lovessey, E. (Ed.), *Contemporary Ergonomics, Proceedings of the Ergonomic Society Annual Conference* (pp. 427-432). London: Taylor and Francis.

Parkes, A. M. (1993). Voice communications in vehicles. In Franzer, S. and Parkes, A. (Eds.), *Driving future vehicles* (pp. 219-228). London: Taylor and Francis.

Parkes, A. M., & Burnett, G. E. (1993). An evaluation of medium range advance information in route-guidance displays for use in vehicles. *Proceedings of the IEEE-IEE vehicle navigation and information systems conference*, 238-241. Piscataway, NJ: Institute of Electrical and Electronic Engineers.

Paul, R. (1996, December). Lost or found? Finding our way through the high-tech terrain of onboard navigation systems. *Motor Trend*, December, 107- 113.

Prabhu, G. V., Shalin, V. L., Drury, C. G., & Helander, M. (1996). Task-map coherence for the design of in-vehicle navigation displays. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, (pp. 882 - 886). Santa Monica, CA: Human Factors and Ergonomics Society.

Redelmeier, D.A. & Tibshirani, R.J. (1997). Assocaition between cellular telephones calls and motor vehicle collisions. *The New England Journal of Medicine,* Vol. 336 (2), pp. 453-458.

Rockwell, T. H (1987). Spare visual capacity in driving -- revisited: New empirical results for an old idea. In A. G. Gale et al. (1988). *Vision in vehicles II* (pp. 317-324). Amsterdam: Elsevier.

Rosen, D. A., Mammano, F. J., and Favout, R. (1970). An electronic route-guidance system for highway vehicles. *IEEE Transactions on Vehicular Technology*, VT-19, 143-152.

Ross, T., Vaughan, G., & Nicolle, C. (1997). Design guidelines for route guidance systems: Development process and an empirical example for timing of guidance instructions. In Y. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces* (pp. 139 - 152). Mawhah, NJ: Lawrence Erlbaum.

Sarno, K., & Wickens, C. (1995) Role of multiple resources in predicting time-sharing efficiency: evaluation of three workload models in a multiple-task setting. The International Journal of Aviation Psychology, 5 (1). 107-130.

Scott, S. (1997). ITS In-vehicle systems safety and human factors standards needs. *Proceedings of the ITS America Seventh Annual Meeting and Exposition.* Washington, DC: Intelligent Transportation Systems of America (CD-ROM).

Serafin, C., Wen, C., Paelke, G., & Green, P. (1993). Car phone usability: a human factors laboratory test. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, pp. 220-224.

Serafin, C., Wen, C., Paelke, G., & Green, P. (1993). *Development and human factors tests of car phones*. Technical Report 93-17. Ann Arbor, MI: University of Michigan Transportation Research Institute.

Shriffen, R. M. (1988). Attention. In R. C. Atkinson, R. J., Herrnstein, and R. D. Luce (Eds.), *Stevens' handbook of experimental psychology*, Second edition (pp. 739-811). New York: John Wiley.

Snowberry, K., Parkinson, S. R., and Sisson, N. (1983). Computer display menus. *Ergonomics*, *26*(7), 699-712.

Society of Automotive Engineers (1999). *SAE Recommended Practice for Navigation and Route Guidance Function Accessibility While Driving* (SAE J2364). Committee Draft submitted for ballot. Warrendale, PA: Society of Automotive Engineers.

Society of Automotive Engineers (1998a). *SAE Recommended Practice for Calculating the Time to Complete In-Vehicle Navigation and Route Guidance Tasks* (SAE J2365). Committee Draft. Warrendale, PA: Society of Automotive Engineers.

Society of Automotive Engineers (1998b). *SAE Recommended Practice for Navigation and Route Guidance Function Accessibility While Driving* (SAE J2364). First distributed committee draft. Warrendale, PA: Society of Automotive Engineers.

Smith, G. L., Jr. (1978). *Work measurement: A systems approach* (pp. 83-98). Columbus, OH: Grid Publishing.

Srinivasan, R., & Jovanis, P. P. (1997). Effect of selected in-vehicle route guidance systems on driver reaction times. *Human Factors*, *39*(2), 200-215.

Stein, A. C., Parseghian, Z., & Allen, R. W. (1987). *A simulator study of the safety implications of cellular mobile phone use*. Hawthorne, CA: Systems Technology, Inc. Paper No. 405 (March). Hawthorne, CA: Systems Technology Inc. (See also 31st Annual Proceedings, American Association for Automotive Medicine, New Orleans, LA, September 1987.)

Streeter, L. A., Vitello, D., & Wonsiewicz, S. A. (1985). How to tell people where to go: comparing navigational aids. *International Journal of Man-Machine Studies*, *22*, 549-562.

Streeter, L. A., and Vitello, D. (1986). A profile of drivers' map-reading abilities. *Human Factors*, *28*(2), 223-239.

Swets, J. A. (1992). The science of choosing the right decision threshold in high-stakes diagnostics. *American Psychologist*, *47*, 522-532.

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Mahwah, NJ: Lawrence Erlbaum.

Tijerina, L., Kiger, Kiger, S., Rockwell, T. H., & Wierwille, W. W. (1995, June). *Final Report Supplement - Task 5 Heavy vehicle driver workload assessment Workload assessment protocol* (Report No. DOT HS 808 417). Washington, D.C.: National Highway Traffic Safety Administration.

Tijerina, L, Kiger, S., Rockwell, T. H., & Tornow, C. E. (1995a). *Final report - Workload assessment of in-cab text message system and cellular phone use by heavy vehicle drivers on the road*. (Contract No. DTNH22-91-07003). DOT HS 808 467 (7A), Washington, DC: U.S. Department of Transportation, NHTSA.

Tijerina, L., Kiger, S. M., Rockwell, T. H., & Tornow, C. (1995b). Workload assessment of in-cab text message system and cellular phone use by heavy vehicle drivers on the road. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting,* pp. 1117 - 1121.

Tijerina, L. (1996, October). *Executive summary: Heavy vehicle driver workload assessment* (Report No. DOT HS 808 466). Washington, DC: National Highway Traffic Safety Administration.

Tijerina, L., Kiger, S. Rockwell, T., and Wierwille, W. (1996, October). *Heavy vehicle driver workload assessment Task 5: Workload assessment protocol* (Report No. DOT HS 808 467). Washington, DC: National Highway Traffic Safety Administration.

Tijerina, L., Parmer, E., and Goodman, M. J. (1998). Driver workload assessment of route guidance system destination entry while driving: A test track study. *Proceedings of the 5th World Congress on Intelligent Transport Systems*, 12-16 October 1998, Seoul, Korea. CD-ROM.

Turnage, J. J., and Kennedy, R. S. (1995). A hierarchical view of human performance: The neglected contribution of visual factors to predicting performance. In A. C. Bittner and P. C. Champney (Eds.), *Advances in industrial ergonomics and safety VII* (pp. 195-202). London: Taylor and Francis.

Van Hoof, K., & Van Strien, J. (1997). Verbal-to-manual and manual-to-verbal dual task interference in left-handed and right-handed adults. *Perceptual and Motor Skills*, *85*, 739-746.

Violanti, J. M., (June, 1997) Cellular phones and traffic accident characteristics, Abstracted in *American Journal of Epidemiology,* Vol. 145 (11), Abstract No. 230.

Violanti, J. M., & Marshall, J.R. (1996). Cellular phones and traffic accidents: an epidemiological approach. *Accident Analysis and Prevention,* Vol. 28, pp. 265-270.

Walker, J., Alicandri, E., Sedney, C., & Roberts, K. (1991). In-vehicle navigation devices: Effects on the safety of driver performance. *Vehicle Navigation and Information Systems Conference Proceedings*, 499-525. Warrendale, PA: Society of Automotive Engineers. Also in *Public Roads*, *56*(1), 9-22.

Waller, P. F., and Green, P. A. (1997). Human factors in transportation. In. G. Salvendy (Ed.), *Handbook of human factors*, Second Edition (pp. 1972 - 2009.

Wang, J., Knipling, R. R. & Goodman, M. J. (1996). The role of driver inattention in crashes; new statistics from the 1995 crashworthiness data system (CDS). *40th Annual Proceedings: Association for the Advancement of Automotive Medicine,* pp. 377-392.

Weimer, J. (1995). Developing a research project. In J. Weimer (Ed.), *Research techniques in human engineering* (pp. 20 - 48). Englewood Cliffs, NJ: Prentice-Hall.

Whitaker, L. A., and Cuqlock-Knopp, V. G. (1995). Human exploration and perception in off-road navigation. In. P. Hancock, J. Flach, J. Caird, and K. Vicente (Eds.), *Local applications of the ecological approach to human-machine systems* (pp. 234-254). Hillsdale, NJ: Erlbaum.

Wiacek, C.J., and Najm, W.G. (1999). *Driver/vehicle characteristics in rear-end precrash scenarios based on the General Estimates System (GES) (Paper No. 99PC-415).* Warrendale, PA: Society of Automotive Engineers.

Wickens, C. D. (1992). *Engineering psychology and human performance* (Second edition). New York: Harper-Collins.

Wickens, C. D., and Carswell, C. M. (1997). Information processing. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 89-129). New York: John Wiley.

Wickens, C., Sandry, D., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, *25* (2). 227-248.

Wierwille, W. W., (1993). Visual and manual demands of in-car controls and displays. In B. Peacock and W. Karwowski (Eds.), *Automotive ergonomics* (pp. 299-320). London: Taylor and Francis.

Wierwille, W., W., & Tijerina, L. (in press). An analysis of driving accident narratives as a means of determining problems caused by in-vehicle visual allocation and visual workload. To appear in A. G. Gale et al. (Eds.), *Vision in Vehicles V*. Amsterdam: North-Holland.

Wierwille, W. W. & Tijerina, L. (1995). *An analysis of driving accident narratives as a means of determining problems caused by in-vehicle visual allocation and visual workload.* Paper presented at the Fifth International Conference on Vision in Vehicles, Glasgow, Scotland, September, 1993. (Conference proceedings, North Holland-Elsevier Press, Amsterdam, 1996.).

Wierwille, W. W., & Tijerina, L. (1995, September). *Modeling the relationship between driver in-vehicle visual demands and accident occurrence*. Paper presented at the Vision In Vehicles VI conference, Darby, England, September, 1995.

Wikman, A., Nieminen, T., & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics*, *41*, 358-372.

Williams, M. J. (1981). Validity of the traffic conflicts technique. *Accident Analysis & Prevention, 13*, 133-145.

Wright, P., Holloway, C., & Aldrich, A. (1974). Attending to visual or auditory verbal information while performing other concurrent tasks. *Quarterly Journal of Experimental Psychology*, *26*, 454-463.

Zaidel, D. M., & Noy, Y. I. (1997). Automatic versus interactive vehicle navigation aids. In Y. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces* (pp. 287-307). Mahwah, NJ: Lawrence Erlbaum..

Zwahlen, H. Adams, C. C., and DeBald, D. P. (1988). Safety aspects of CRT touch panel controls in automobiles. In A. G. Gale, et al. (Eds.), *Vision in Vehicles II* (pp. 335-344). London: Taylor and Francis.

Zwahlen, H. T., Adams, Jr. C. C., & Schwartz, P. J. (1988). Safety aspects of cellular telephones in automobiles (Paper No. 88058). *Proceedings of the ISATA Conference*, Florence, Italy.

Appendix A: Informed Consent Form for the Destination Entry Study

TEST PARTICIPANT CONSENT FORM

Title of Study: : **Route Navigation System Data Entry Study No. 1**

Study Description: Route navigation systems are being developed and marketed for use in cars and trucks. Such systems allow a driver to enter a destination into an in-vehicle computer and receive visual or verbal instructions on how to get there. These devices introduce additional tasks which might compete with the driver's primary job of safely controlling the vehicle at all times. The National Highway Traffic Safety Administration (NHTSA) is conducting research to measure the effects on drivers of introducing such high-technology devices. One area of research is the effects on driver behavior and performance of entering desired destinations into route navigation systems while driving (hereafter referred to as destination entry).

The purpose of this study is to gather data on driver behavior and performance while attempting to enter destinations into various route navigation systems while driving. As a participant, you will receive training on how to complete destination entry tasks with each of several commercially available route guidance systems. You will have an opportunity to practice entering destinations with each system in a parked vehicle and ask any questions you wish of the experimenter. The experimenter will then show you a set of "target" destinations that you are to write down on individual 4x6 index cards so that you will be able to use them for destination entry while driving on the TRC test track. You will be asked to drive a series of laps around the TRC test track. At various points during these laps, the ride-along experimenter will ask you to maintain certain stated speeds. The ride-along experimenter will also periodically hand you a 4x6 index card with a destination written on it, a signal that you are to enter that destination into a specific route guidance system, WHEN AND IF YOU BELIEVE IT IS SAFE TO DO SO GIVEN THE CURRENT DRIVING CONDITIONS ON THE TRACK. After you have completed a series of laps, you will return to Building 60 and will answer a series of debriefing questions. At the end of all laps and the debriefings, the test session will be completed.

It is very important to always remember that you, as the driver, are in control of the vehicle and you must be the final judge on when or whether to respond to any request. You should follow a request or complete a maneuver only when, in your judgement, it is safe and convenient to do so. The ride-along experimenter will not be able to insure safety; you as the driver are responsible for that.

Remember, safety while driving on the test track is your primary responsibility. Complete requests only when and if you believe it is safe to do so.

Risks: While driving for this study, you will be subject to all risks normally associated with driving on the TRC test track plus any additional risks associated with completing in-vehicle tasks while driving. There are no known physical or psychological risks associated with participation in this study beyond those indicated.

Be aware that accidents can happen any time when driving. You remain responsible for your driving during this testing. If the ride-along experimenter should make a request, you are not to do it unless you judge it is safe to do so.

Benefits: This testing will provide data on driver behavior, performance and, judgements regarding destination entry with the selected route guidance systems while driving. This data will provide a scientific basis for guiding recommendations on standards for route guidance systems in the future.

Confidentiality: The data recorded on you will be analyzed along with data gathered from other test participants during this testing. Your name will not be associated with any final report, publication, or other media that might arise from this study. However, your video-taped likeness (in video-tape or still photo formats created from the video-tape) and engineering data from you specifically may be used for educational and research purposes. A waiver of confidentiality for permission to use the video-tape and engineering data (including data or images derived from these sources) is included for you to sign as part of this form. It is not anticipated that you will be informed of the results of this test.
Informed Consent: By signing below, you agree that participation is voluntary and you understand and accept all terms of this agreement. **You have the option of not performing any requested task at any time during the test without penalty.**

Compensation: Should you agree to participate in this testing, it will be considered part of your normal work day activities. There is no special compensation associated with participation in the test.

Principal Investigator: Contact Dr. Louis Tijerina (TRC) or Dr. Riley Garrott (NHTSA VRTC) if you have questions or comments regarding this study. They may be reached at the address and phone number given below

Vehicle Research and Test Center
10820 SR 347
East Liberty, OH 43319
Phone: (937) 666-4511

Disposition of Informed Consent: The VRTC will retain a signed copy of this Informed Consent form. A copy of this form will also be provided to you upon completion of participation in the study.

INFORMED CONSENT:

I, _____, UNDERSTAND THE TERMS OF THIS AGREEMENT AND VOLUNTARILY CONSENT TO PARTICIPATE.

_____ _____
Signature Date

_____ _____
Witness Date

WAIVER OF CONFIDENTIALITY:

I, _____, grant permission, in perpetuity, to the National Highway Traffic Safety Administration (NHTSA) to use, publish, or otherwise disseminate the video-tape (including still photo formats derived from the videotape) and engineering data collected about me in this study for educational, outreach, and research purposes. I understand that such use may involve widespread distribution to the public and may involve dissemination of my likeness in videotape or still photo formats, but will not result in release of my name or other identifying personal information.

_____ _____

Signature Date

_____ _____

Witness Date

Appendix B: Informed Consent Form for SAE J2364 15-Second Rule Study

TEST PARTICIPANT CONSENT FORM

Title of Study: : **Route Navigation System Data Entry Study No. 2**

Study Description: Route navigation systems are being developed and marketed for use in cars and trucks. Such systems allow a driver to enter a destination into an in-vehicle computer and receive visual or verbal instructions on how to get there. These devices introduce additional tasks which might compete with the driver's primary job of safely controlling the vehicle at all times. The National Highway Traffic Safety Administration (NHTSA) is conducting research to measure the effects on drivers of introducing such high-technology devices. One area of research is the effects on driver behavior and performance of entering desired destinations into route navigation systems while driving (hereafter referred to as destination entry).

The purpose of this study is to gather data on driver behavior and performance while attempting to enter destinations into various route navigation systems both while driving the vehicle and when it is stationary. You will also be asked to complete several additional tasks for comparison purposes. These will include dialing a cellular telephone, tuning a car radio, and adjusting the heating, ventilation, and air conditioning (HVAC) system. As a participant, you will receive training on how to complete destination entry tasks with each of several commercially available route guidance systems. The experimenter will show you a set of "target" destinations that you are to write down on individual 4x6 index cards so that you will be able to use them for destination entry during practice and static testing as well as while driving on the TRC test track. You will have an opportunity to practice entering destinations with each system in a parked vehicle and ask any questions you wish of the experimenter. You will be asked to drive a series of laps around the TRC test track. At various points during these laps, the ride-along experimenter will ask you to maintain certain stated speeds. The ride-along experimenter will also periodically hand you a 4x6 index card with a destination written on it, a signal that you are to enter that destination into a specific route guidance system, WHEN AND IF YOU BELIEVE IT IS SAFE TO DO SO GIVEN THE CURRENT DRIVING CONDITIONS ON THE TRACK. After you have completed a series of laps, you will return to Building 60. Before or after this testing, the experimenter will ask you to perform a similar test while the vehicle is parked in a work bay. The experimenter will hand you a 4x6 index card with a destination written on it, a signal that you are to enter that destination into a specific route guidance system while the vehicle is stationary. A similar procedure will be followed for the comparison tasks. At the end of all laps and static testing, the test session will be completed.

It is very important to always remember that you, as the driver, are in control of the vehicle and you must be the final judge on when or whether to respond to any request or engage in any task. You should follow a request or complete a task or complete a maneuver only when, in your judgement, it is safe and convenient to do so. The ride-along experimenter will not be able to insure safety; you as the driver are responsible for that.

Remember, safety while driving on the test track is your primary responsibility. Complete requests only when and if you believe it is safe to do so.

Risks: While driving for this study, you will be subject to all risks normally associated with driving on the TRC test track plus any additional risks associated with completing in-vehicle tasks while driving. There are no known physical or psychological risks associated with participation in this study beyond those indicated.

Be aware that crashes can happen any time when driving. You remain responsible for your driving during this testing. If the ride-along experimenter should make a request, you are not to do it unless you judge it is safe to do so.

Benefits: This testing will provide data on driver behavior, performance and, judgements regarding destination entry with the selected route guidance systems both in static mode and while driving. This data will provide a scientific basis for guiding recommendations on standards for route guidance systems in the future.
Confidentiality: The data recorded on you will be analyzed along with data gathered from other test participants during this testing. Your name will not be associated with any final report, publication, or other media that might arise from this study. However, data from you specifically may be used for educational and research purposes. A waiver of

confidentiality for permission to use the data (including data or images derived from these sources) is included for you to sign as part of this form. It is not anticipated that you will be informed of the results of this test.

Informed Consent: By signing below, you agree that participation is voluntary and you understand and accept all terms of this agreement. **You have the option of not performing any requested task at any time during the test without penalty.**

Compensation: Should you agree to participate in this testing, it will be considered part of your normal work day activities. There is no special compensation associated with participation in the test.

Principal Investigator: Contact Dr. Louis Tijerina (TRC) or Dr. Riley Garrott (NHTSA VRTC) if you have questions or comments regarding this study. They may be reached at the address and phone number given below

Vehicle Research and Test Center
10820 SR 347
East Liberty, OH 43319
Phone: (937) 666-4511

Disposition of Informed Consent: The VRTC will retain a signed copy of this Informed Consent form. A copy of this form will also be provided to you.

INFORMED CONSENT:

I, _____, UNDERSTAND THE TERMS OF THIS AGREEMENT AND VOLUNTARILY CONSENT TO PARTICIPATE.

_____ _____
Signature Date

_____ _____
Witness Date

WAIVER OF CONFIDENTIALITY:

I, _____, grant permission, in perpetuity, to the National Highway Traffic Safety Administration (NHTSA) to use, publish, or otherwise disseminate the data collected about me in this study for educational, outreach, and research purposes. I understand that such use may involve widespread distribution to the public, but will not result in release of my name or other identifying personal information.

_____ _____
Signature Date

_____ _____
Witness Date