Technical Report UMTRI-97-19 January, 1998

Usability of Menu-Based Interfaces for Motor Vehicle Secondary Functions

Minoru Sumie,
Charmian Li,
and Paul Green

UMTRI HUMAN FACTORS

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTRI-97-19		

Usability of Menu-Based Interfaces for Motor Vehicle Secondary Functions	5. Report Date January, 1998 6. Performing Organization Code Account 391551
7. Author(s) Minoru Sumie, Charmian Li, and Paul Green	8. Performing Organization Report No. UMTRI-97-19
9. Performing Organization Name and Address The University of Michigan Transportation Research Institute 2901 Baxter Rd, Ann Arbor, Michigan 48109-2150 USA	10. Work Unit no. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Mitsubishi Motors Corporation 1, Nakashinkiri, Hashime-cho, Okazaki, Japan 444	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

minor corrections made, June 1998

16. Abstract

Three experiments were conducted to develop methods and data for examining the usability of secondary controls for accessing hierarchical menus system while driving. Hierarchical menu interfaces may be used in future vehicles for the operation of radio and navigation systems.

In the first experiment, eight drivers performed a secondary task in an instrumented car and the UMTRI driving simulator. The real and simulated roads were not matched. There were no differences in secondary task time-related measures between contexts, but absolute validity of the primary task was not high. As in previous studies, speed variance was lower and lane variance was greater in the simulator. This experiment established that the UMTRI simulator was a reasonable context for examining the use of secondary controls.

In the second experiment, two subjects drove the simulator for brief periods of time over 20 successive workdays. Secondary task performance continued to improve across the entire period, with most of the performance gains being evident by the fifth day. There were few interactions of time with test conditions, suggesting that testing across a large number of days was not necessary.

In the third experiment, 16 subjects drove the simulator while retrieving information using a menu system. In terms of task completion times, the number keypad yielded the best performance, with the rotary knob and the trackball tied for second, and the touchpad yielding the worst performance. As for the menu structures, the order from best to worst was 8 x 2 (8 items, 2 levels deep), 4 x 3, and 2 x 6, though the differences between 8 x 2 and 4 x 3 were small.

17. Key Words	18. Distribution Statement
control and display, compatibility, ITS, human factors, safety, simulation, instrumented vehicles	No restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
None	None	99	

Form DOT F 1700 7 (8-72) **Reproduction of completed page authorized**

Usability of Menu-Based Interfaces for Motor Vehicle Secondary Functions

Minoru Sumie, Charmian Li, and Paul Green

UMTRI Technical Report 97-19, January 1998

University of Michigan, Ann Arbor, Michigan, USA

1 PURPOSE

1. Assess validity of simulator for determining usability of secondary controls
2. Develop simulator-based test protocol
3. Compare ease of use of alternative controls (number pade, knob, trackball, touchpad) and menu structures (8 wide x 2 deep, 4 x 3, 2 x 6)

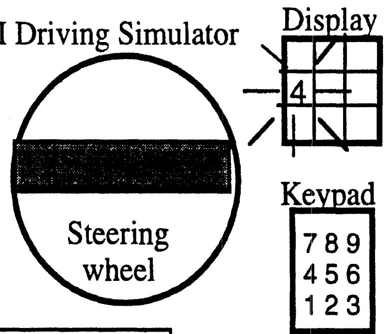
2 EXPERIMENT 1 - SIMULATOR VALIDATION

METHOD

ISSUE Are primary task, secondary task, and eye fixation measures collected in the UMTRI simulator indicative of those collected on the road in an instrumented car? Can the simulator be used in place of on-the-road testing?

Dependent Measure
Interbutton response time
Task completion time
IP glance duration
of IP glances
Forward-viewing time
SD of steering angle
SD of standardized lane pos
SD of throttle position
SD of vehicle speed

UMTRI Driving Simulator



Where	Road driven	Mi/hr Traffic	# Button presses
On-Road	1. rural road	40 light	25
	2. residential road	30 light	25
	3. suburban street	45 moderate	25
	4. expressway	55 moderate	25
UMTRI Simulator	5. practice	none	no 2nd task
	6. winding road	35 none	25
	7. winding road	50 none	25
	8. winding road	75 none	25

SUBJECTS

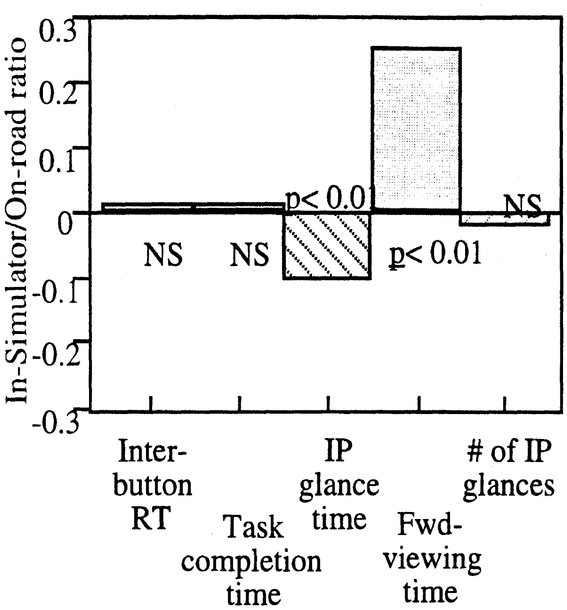
Age	♂	♀
<30	2	2
>65	2	2

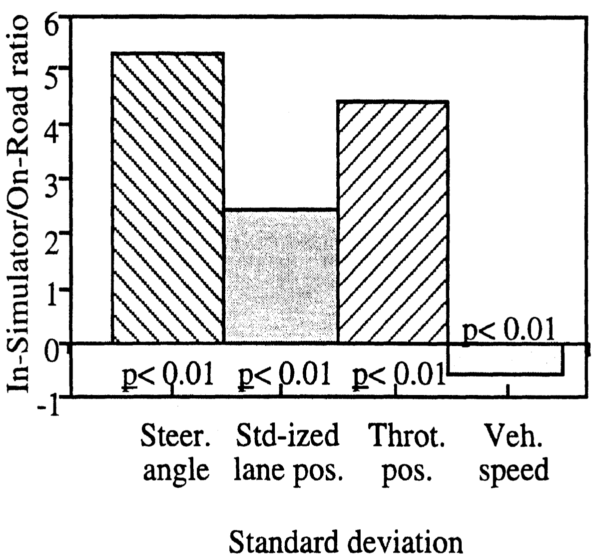
RESULTS

ANOVA Comparison of Simulator and On-Road Data

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++	++	++	++	++
Gender (G)	++	++	++	++	++			++	++
Subject (S)	++	++	++	++	++			++	++
Context (C)			++	++		++	++	++	++
Task (T)	n/a	n/a	n/a	n/a	n/a	++	++	++	++
A x G	++	++	++	+				+	++
A x C			++	+		+	++	++	++
A x T	n/a	n/a	n/a	n/a	n/a	++			
G x C								+	++
G x T	n/a	n/a	n/a	n/a	n/a				+
S x C	++		++	++	+		+	++	++
S x T	n/a	n/a	n/a	n/a	n/a				
C x T	n/a	n/a	n/a	n/a	n/a		++	++	

(++: p< 0.01; +: p< 0.05), n/a=not available





CONCLUSIONS

- There were few differences in secondary task times.
- The driving data was more variable in the simulator. (However roads were not matched and simulated roads were more difficult to drive.)
- The use of the simulator was acceptable for secondary control studies.

3 EXPERIMENT 2 - EFFECT OF PRACTICE ON DRIVER PERFORMANCE IN THE SIMULATOR

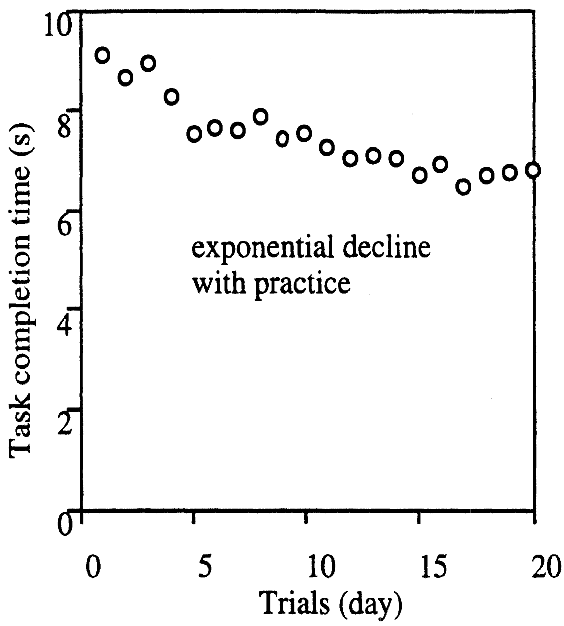
ISSUE Is data from a single session in a driving simulator adequate to examine secondary control usability? Are there interactions of performance with practice?

METHOD

Two young subjects drove the simulator for a short period of time for 20 days, A winding road from Experiment 1 was driven. There were 3 test blocks/day.

Block 1	Block 2	Block 3
35 mi/hr 25 button presses	50 mi/hr none	50 mi/hr 25 button presses

RESULTS



ANOVA Simulator Learning Measures

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Subject (S)	++	++	n/a	n/a	n/a	++	++		
Condition(C)			n/a	n/a	n/a	++	++		
Day (D)	++	++	n/a	n/a	n/a	++	++	++	
S x C			n/a	n/a	n/a	++	++		
S x D	++	++	n/a	n/a	n/a			+	+
C x D			n/a	n/a	n/a	+			

(++: p< 0.01; +: p< 0.05), n/a = not available

CONCLUSION There were no interactions between conditions and day and practice effects can be modeled as an exponential process, so 1 session is enough.

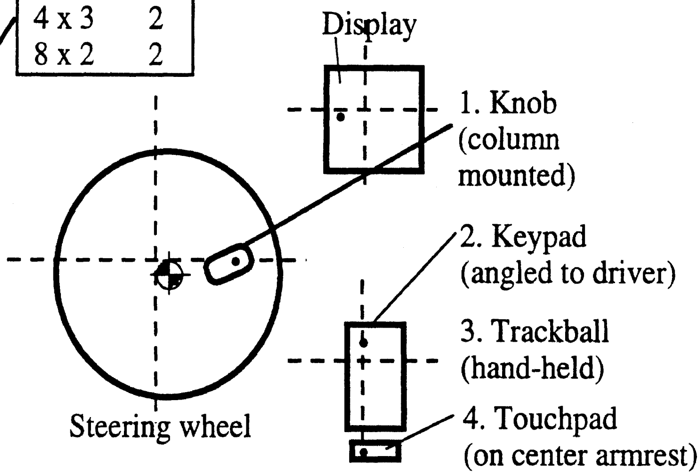
4 EXPERIMENT 3 - USABILITY OF HIERARCHICAL MENU INTERFACES

ISSUE While driving, how easy to use were various controls (number pad, knob, trackball, touchpad) for hierarchical menu systems as a function of menu depth and breadth?

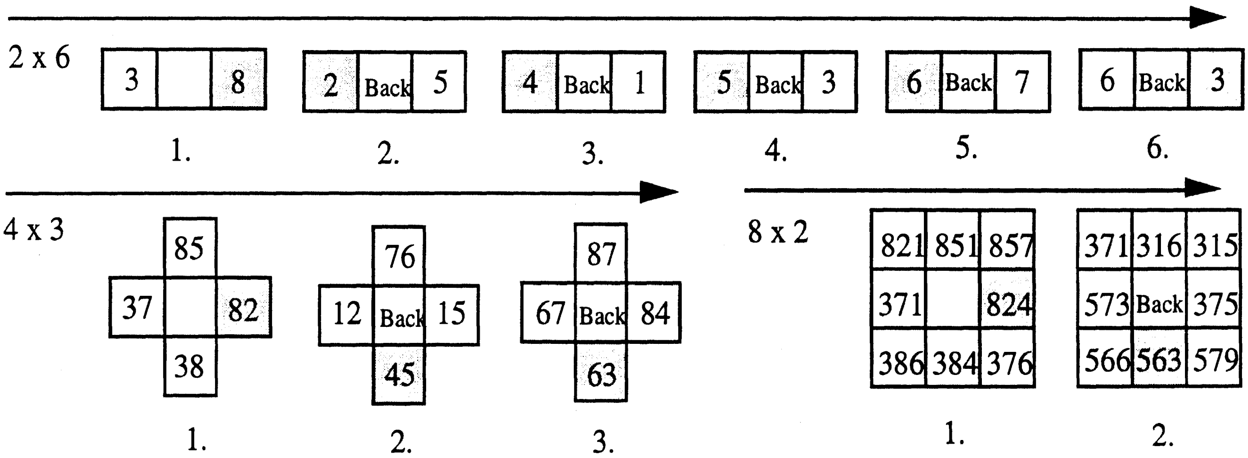
METHOD

		Trials/Menu
Practice	number pad	3
	knob	3
	trackball	3
	touchpad	3
	none-no 2nd task	(0)
Test -35 mi/hr	number pad	6
	knob	6
	trackball	6
	touchpad	6
Test -50 mi/hr	number pad	6
	knob	6
	trackball	6
	touchpad	6

Menu	Trials
2 x 6	2
4 x 3	2
8 x 2	2



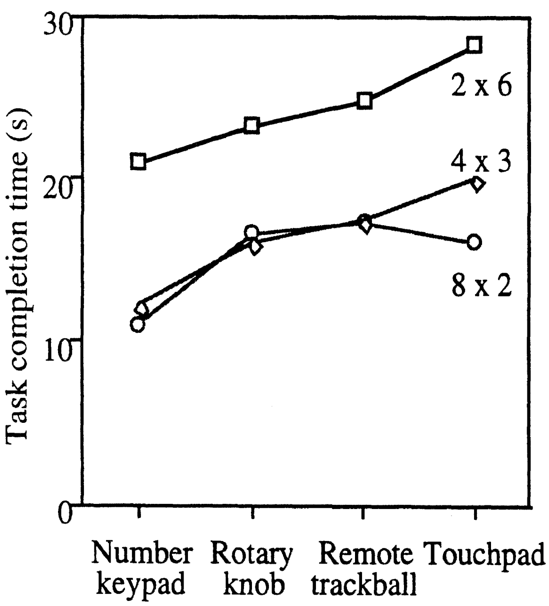
How "824563" is entered for each of the 3 menu structures.

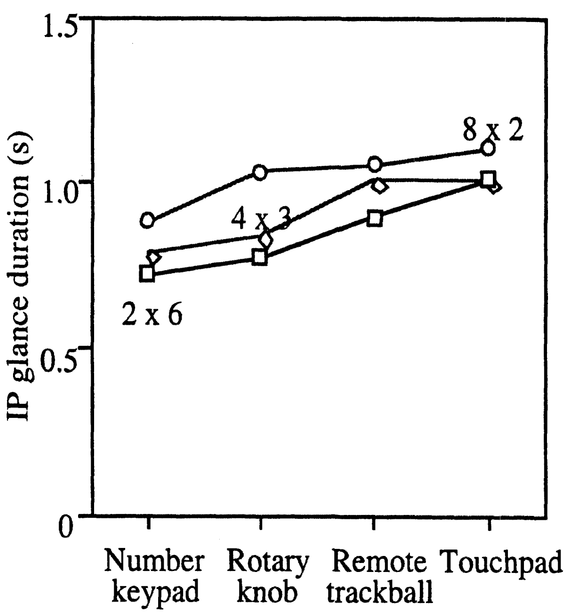


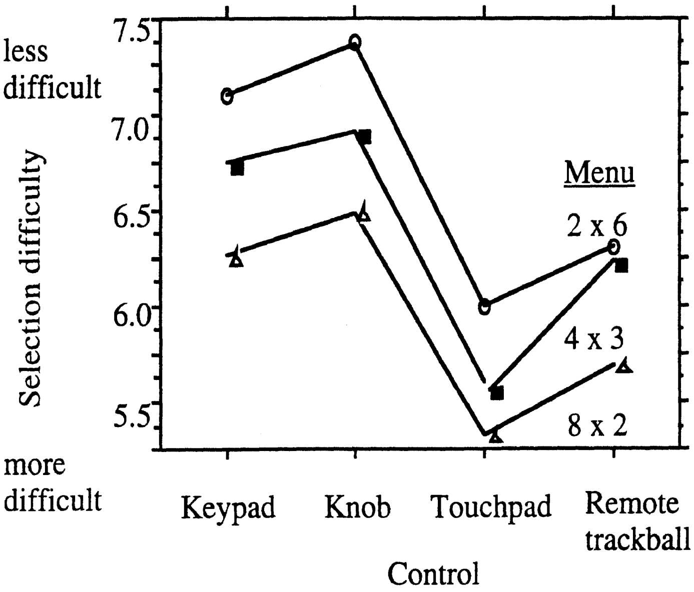
RESULTS ANOVA of Data from Experiment 3

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++	++	++	++	++
Gender (G)		++	+	++					++
Subject (S)	++	++	++	++	++	++	++	++	++
Speed (Sp)	++	++	++	+		++		++	++
Control (C)	++	++	++		++	++			+
Menu (M)	++	++	++	++	++	n/a	n/a	n/a	n/a
A x G					+	++		++	
A x Sp			++						
A x C		+	++						
A x M						n/a	n/a	n/a	n/a
G x Sp			++					++	++
G x C	++			++					
G x M						n/a	n/a	n/a	n/a
S x Sp		++	++	+		+			
S x C		++	++	++	++	+			
S x M		+				n/a	n/a	n/a	n/a
Sp x C			++						
Sp x M						n/a	n/a	n/a	n/a
C x M	++	++	+	+		n/a	n/a	n/a	n/a

(++: p< 0.01; +: p< 0.05)







Ratings - Easy to ...

	Access	Operate	Drive
Knob	7.9	7.6	6.5
Trackball	7.7	6.9	6.3
Keypad	6.6	7.7	6.1
Touchpad	6.6	5.5	5.0

1=difficult
10=not difficult

CONCLUSIONS:

1. A keypad is the best choice for a control and a knob is a second choice for the following reasons:

   A keypad
   - had the smallest task completion times (1/3 less than knob),
   - have relatively better performance as the number of menu items grew (Hick's Law result),
   - needed fewer and shorter glances to use
   - was rated as easiest to operate and most accessible
   - led to the same driving performance as other controls.

   The recommendation depends on mounting location and other implementation factors. (So in practice, at times, other controls may be better.)

2. There is no clear recommendation for a desired menu structure.
   - Interbutton RTs were least for 2 x 6 menus (about 4 s) (vs. 5 s for 4 x 3, and 7 s for 8 x 2.
   - Task completion times were 15 s for 4 x 3 and 8 x 2 menus, 25 s for 2 x 6.
   - Glance durations were 1.0 s to 8 x 2, under 0.9 s for 4 x 3, and 0.8s for 2 x 6 but 8 x 2 had fewest glances (7) followed by 4 x 3 (7.5) and 2 x 6 (11).
   - The 2 x 6 was easiest to select menu structure for selecting one item, but that may not be important.
   - The driving data did not distinguish among menus, but the results were confounded with practice.

3. In future work, more baseline data is needed.

4. In future work, HUDs and steering wheel controls should be examined.

ACKNOWLEDGMENTS

This research was accomplished as a second collaborative project between the Human Factors Division of the University of Michigan Transportation Research Institute (UMTRI) and Mitsubishi Motors Corporation. The goal of this collaboration was to collect information useful for the development of safer and easier to use driver interfaces for motor vehicles of the future. Previous research concerned driver eye fixations (Hada, 1994). This second collaboration was conducted to examine driver use of secondary controls during a two-year visit of Minoru Sumie (a Mitsubishi engineer) to UMTRI.

The authors wish to acknowledge the assistance of several members of the Human Factors Division at UMTRI:

Aaron Steinfeld	for the preparation of the controls and display, and for simulator experiment setup;
Stewart Katz	for instrumented test vehicle set-up;
Sokha Chau	for eye fixation and driving performance data analysis;
Tammy Rice	for eye fixation and driving performance data analysis;
Edgar Manalo	for eye fixation data analysis; and
Eric Traube	for comments on subject instructions.

The authors also wish to thank members of the Human Factors Group at Mitsubishi Motors for their ideas.

Hideki Hada	for comments on experimental design, data analysis and
Toru Hara	for comments on data analysis.

In addition, the authors wish to thank Mr. Koichi Kamiya and Mr. Yoshihiro Goi, the manager and former manager of the Human Factors group at Mitsubishi Motors, for their guidance of the second collaboration between the two organizations.

TABLE OF CONTENTS

1. INTRODUCTION 1
1.1 SPEECH INPUT AND OUTPUT 1
1.2 HIERARCHICAL MENU SYSTEMS 2
1.3 OVERVIEW OF EXPERIMENTS CONDUCTED 6
2. EXPERIMENT 1 - VALIDATION OF A DRIVING SIMULATOR FOR SECONDARY TASK EVALUATION 7
2.1 BACKGROUND AND ISSUES 7
2.2 TEST PLAN 8
2.2.1 Subjects 8
2.2.2 Roads Driven 8
2.2.3 Instrumented Car 10
2.2.4 Driving Simulator 12
2.2.5 In-Vehicle Secondary Task 14
2.2.6 Procedure Overview 15
2.2.7 Experimental Design 16
2.3 RESULTS 17
2.3.1 Secondary Task Performance (Combined Comparison) 18
2.3.2 Driving Performance (Combined Comparison) 21
2.3.3 Secondary Task Performance (Separate Analyses) 24
2.3.4 Driving Performance (Separate Analyses) 26
2.4 DISCUSSION 30
3. EXPERIMENT 2 - EFFECT OF PRACTICE ON DRIVER PERFORMANCE IN THE SIMULATOR 33
3.1 INTRODUCTION 33
3.2 METHODS 34
3.2.1 Subjects 34
3.2.2 Driving Simulator and Test Route 34
3.2.3 Experimental Design and Procedure 34
3.3 RESULTS 35
3.3.1 Secondary Task Performance 35
3.3.2 Driving Performance 37
3.4 DISCUSSION 38
4. EXPERIMENT 3 - USABILITY OF HIERARCHICAL MENU INTERFACES 39
4.1 INTRODUCTION 39
4.2 METHODS 39
4.2.1 Subjects 39
4.2.2 Test Route 40
4.2.3 Equipment and Materials 40
4.2.4 Hierarchical Menu Structures 41
4.3 RESULTS 46
4.3.1 Secondary Task Performance 46
4.3.2 Driving Performance 53
4.3.3 Post-Trial Subjective Ratings of Menu Usability 54
4.3.4 Post-Trial Subjective Ratings of Control Usability 58
4.4 DISCUSSION 60
REFERENCES 63
APPENDIX A - CONSENT FORM 67
APPENDIX B - BIOGRAPHICAL FORM 69
APPENDIX C - DIFFICULTY RATING FORM 71
APPENDIX D - DATA SUMMARY 75
APPENDIX E - ANOVA TABLES 79

1. INTRODUCTION

The development of electronics has changed the motor vehicle. Electronic controls have helped reduce vehicle-related pollution (through the use of precision fuel injection), enhanced driver safety (through anti-lock braking systems and air bags), and made driving easier and more comfortable (through improved power steering and automatic transmission systems). The next major set of developments in improving motor vehicles is associated with Intelligent Transportation Systems (ITS). Features to be implemented include navigation, traffic information, and collision avoidance systems. Implementations of some of these features have already occurred, especially in Japan. Each of these systems is intended to interact with the driver, with controls for the driver to use and displays for the driver to see or hear. The proliferation of controls and displays poses a concern that both the new systems and the expanded functionality of existing systems could overwhelm drivers. Of concern are training requirements (you need to take a class to learn how to operate this car), customer satisfaction (but I know how to operate a radio with five buttons), and driver safety.

Safety could be compromised if drivers are diverted from the primary task of driving by operating and responding to either new (e.g., navigation) or more complex existing systems (e.g., audio, climate control). Driver inattention is widely recognized as a major cause of traffic accidents. Statistics in NHTSA Traffic Safety Facts (1994) indicated that 6 percent of all fatal crashes were caused by inattentiveness of drivers. When inattention is broadly defined (to include improper lookout, misperception and distraction), other studies have estimated 25 to 50 percent of all crashes are due to inattention (Zaidel, Paarlberg, Shinar, 1978). Thus, although there is debate as to what inattention is, how it is measured, and the validity of the statistics in the literature, inattention is nonetheless an important issue.

Driving is primarily a visual task (Sivak, 1996). While driving and concurrently performing a secondary task, drivers alternately look at the road scene and the vehicle interior (instrument panel, center console, etc.), where the secondary task is located. The need to share visual input and central processing resources between tasks can degrade one or both tasks. Hence, to assure safe driving, it is necessary to minimize the visual and attentional demands from in-vehicle secondary systems and the intrusiveness of secondary tasks on driving.

Several alternatives to the existing approach of dedicated controls and displays have been proposed to overcome potential driver overload. Two leading proposals are: (1) the use of voice (primarily for presenting information to the driver) and (2) hierarchical menus with integrated controls.

1.1 SPEECH INPUT AND OUTPUT

Speech output to the driver and in-vehicle speech recognition systems could reduce the driver's visual demands, eliminating the need to see displays and visually guide their hands to controls. As an example, many contemporary route guidance systems provide turn-by-turn voice guidance, and some systems sold in Japan respond to commands spoken by the driver. The supporting rationale for this approach is multiple-resource theory (Wickens, 1992). The theory proposes that time-sharing is more efficient when the two tasks use different modalities. Thus, in theory, drivers should find performing a spatial and a verbal task (looking and listening) in parallel easier than performing two spatial tasks (both looking). However, the auditory input channel is quite limited in bandwidth and supports only serial processing. Vision has a high bandwidth and allows parallel processing. Thus, in some situations a driver may be able to get more information from a sampled visual channel than a continuous auditory channel.

A second key distinction is that auditory information must be remembered, whereas visual information from secondary controls (in most cases) is always available, provided the driver can look at the display.

A third consideration is instrument panel real estate. In brief, speech input and output requires essentially no additional front panel space, since output is typically via existing speakers. Input is via a microphone, often mounted in unused space in the A-pillar, or if mounted elsewhere, requiring a minimum of space. On the other hand, manual controls always occupy some additional space within the vehicle, be it a module on the center console or controls on a stalk mounted on the steering column.

1.2 HIERARCHICAL MENU SYSTEMS

In conventional in-vehicle secondary systems, controls and displays are grouped together (e.g., audio, climate control, navigation, etc.). To make space for additional controls and displays, one could reduce the size of existing controls and displays. Such an approach would have major negative consequences, such as labels too small to read, controls too small to grasp, or information in the wrong format (e.g., digital information where a scale is desired). The alternative is therefore to consider sharing some controls and displays. In that case, some means is needed to identify the function served by a control and the information presented by a display. If the functionality of the control provided is extensive, a collection of menus, usually hierarchically organized, may be needed. An early description of this idea appears in Green (1979).

While examples of hierarchical menu systems are quite common for navigation systems, there are few well known examples that exhibit broader functionality. Figure 1.1 shows the menu structure of the Rockwell PathMaster (also sold by Zexel, Oldsmobile, and Siemens).

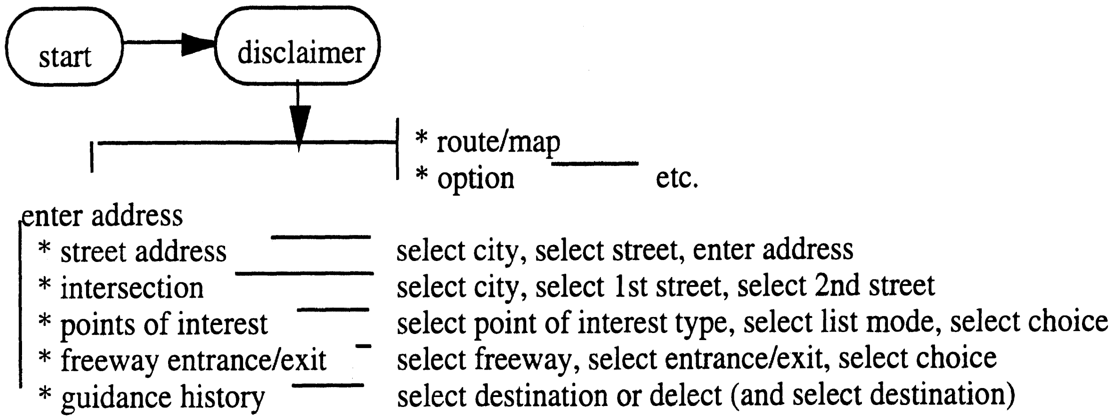


Figure 1.1. PathMaster driver interface logic

One example of an instrument cluster design employing hierarchical menu structures is the Delco Electronics Eyes Forward product (Figure 1.2). (See Heuchert, 1995). All controls were integrated into a single interface. Four steering wheel-mounted rocker switches activated the reconfiguration of the flat panel LCD in the speedometer/odometer cluster. The instrument cluster displayed labels for the buttons on the steering wheel, labels that changed depending on the position within the menu hierarchy. Consequently, drivers could operate the controls without moving their hands from the steering wheel, eliminating the need for a glance to guide the movement of their hands to a control. This interface style may be the trend for in-vehicle secondary systems in future cars.

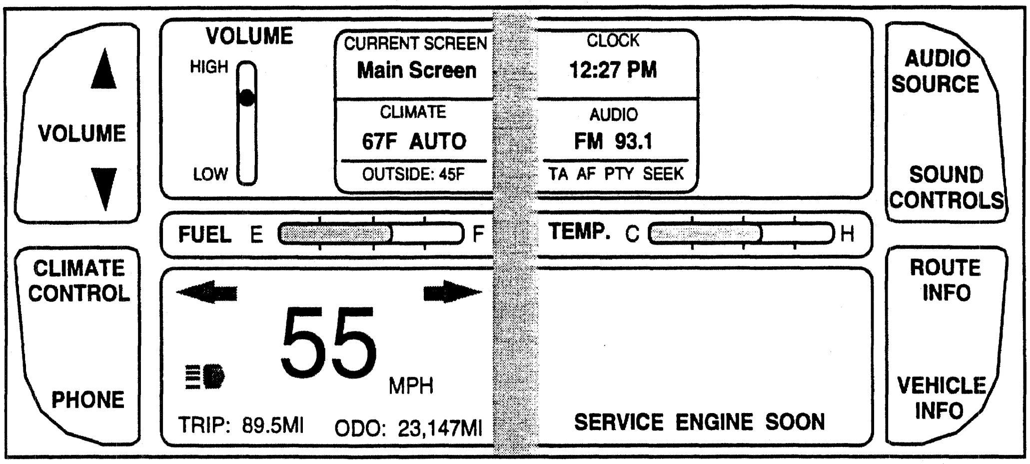


Figure 1.2. Delco Eyes Forward driver interface
(four centrally grouped displays, reconfigurable switch labels appear at the edges of the figure)

The most critical requirement for a display unit is its visibility. To minimize eyes-off-the-road time, the display unit should be located as high on the IP (instrument panel) as possible so that it is close to the driver's line of sight to minimize eye travel. The focal distance (from the driver to the display) should be as far away as possible to minimize the need to reaccommodate from the road. Over the past decade, car manufacturers have introduced Head-Up Displays (HUDs) into production vehicles, displaying limited information (speed, turn signals, warning symbols, etc.) on the windshield. Previous studies suggested that such displays improved driver's ability to detect hazardous situations in the forward scene (Kiefer and Gellatly, 1996). Although it is certain that HUDs contribute to safer driving, the high production cost has limited customer demand for this option. Further, the primary information element on contemporary HUDs is the speed; however, glances to the speedometer are sufficiently infrequent that the aggregate effect of reducing total speedometer glance time on the total time for all driving-related glances is minimal.

The most important consideration for a control is its accessibility. The controls should be located as close to the driver as possible (within easy reach distance), and should be visible to the driver, especially if the driver is unaccustomed to the controls. However, satisfying both the visibility and accessibility constraints can be a challenge. For instance, the center console on the IP where conventional in-vehicle controls are located has fairly good visibility, but the reach distances for drivers can be large. The steering stalk and the steering pad are better choices for accessibility to minimize hand-reach distance, but they offer poor visibility, especially for older drivers who struggle to focus on close objects. Furthermore, stalk controls may be obstructed by the steering wheel.

One potential solution to the control location dilemma (locating the control high and far away optimizes legibility, but placing it low and close to facilitates access) is to make the control so easy to locate and selections so apparent from the display that minimal visual guidance is needed, especially after practice. This is possible using steering wheel-mounted controls and a remote display, as is the case for some hierarchical menu systems.

There has been a fair amount of research on hierarchical menu systems. Miller (1980, 1981) presents a very straightforward analysis of hierarchical menus basing his approach on the work of Posner. Posner suggested that matching decisions can be categorized into three groups as suggested in the following table. The decision times listed are from various response time experiments in the literature (reviewed by Miller) involving common words and represent theoretical times for menu selection.

Table 1.1. Posner's data on matching

Match type	Example of "same" match	Decision time (s)
Physical	AA (same form of letter A)	
Identity	Aa (both letter A)	.40 + .08 (# choices)
Category	Ab (both letters)	.40 + .17 (# choices)

Most menu selections involve (1) determining which category an item is in for top level menus (category match) and (2) matching the final item (an identity match). Thus, in theory, the time to select an item from a hierarchical menu should be equal to the times for all the category matches plus the time for the final identity match.

In Miller's experiment, subjects were shown the goal item (e.g., horse), and then instructed to navigate through a series of menus until the matching term was located. At each level there were either 2, 4, 8, or 64 choices, but the number of end nodes, 64, remained constant. The results showed that a menu hierarchy of two levels had the fastest task completion time, produced the fewest number of errors, showed the least amount of variability, and was the easiest to learn. Therefore, the recommendation for display design was that an expansion in breadth was better than an expansion in depth.

Based on his experiment, Miller proposed the equations shown in Table 1.2 to predict performance. The differences in the identity match time parameters proposed by Miller and those based on the literature (Table 1.1) were small. While the parameters for category match times were somewhat different, the predicted times for short menus (3-5 items) were fairly close. For larger number of choices, some of the differences in category match time are offset by differences in identity match time. One must bear in mind that these data are from two different groups of people completing different protocols.

Table 1.2. Decision times from Miller's data

Match type	Decision time (s)
Identity	.32 + .11 (# choices)
Category	.80 + .035 (# choices)

In summary, the key elements of Miller's approach were that (1) the decision process had two aspects to it: identity and category, (2) decision times were a linear function of the number of choices, and (3) Miller's results agreed with the theoretical data of Posner.

A second study was done by Landauer and Nachbar (1985). The study examined menu selection from a touch screen, where up to 4096 items could be presented. Landauer and Nachbar proported that menu selection should fit the well known Hick-Hyman expression in the literature, where decision time is proportional to $\log_2$ N (N being the number of choices). Averaging across levels of practice, Landauer and Nachbar reported the following results shown in Table 1.3. These times are somewhat larger than those reported elsewhere. The differences between experiments are most likely due to different presentation and response collection methods.

Table 1.3. Decision times to select one item from a screen

Item	Decision time (s)
Number	0.77 + .73 (log2 (# choices))
Word	1.25 + .56 (log2 (# choices))

While these predictions are somewhat different from those proposed by Miller, the larger constants probably reflected a more difficult decision task. Furthermore, when the number of choices was small, so too was $\log_2$ of the number of choices.

As an example of more recent work, Musseler (1994) had subjects find names in a hierarchical menu. In the first experiment, subjects were instructed to find matching items using a graphical user interface (GUI) and a mouse. In the second experiment, subjects used a control key paired with various letter keys to make a selection.

The menu bar had six items. Each item on the menu bar was the title for a drop-down submenu of two to seven items. Submenus were drop down (activated when the cursor entered the field), not pull down (requiring the user to hold down the mouse button and position the cursor to where the submenu was to be selected). Three sets of variables were considered (Table 1.4). (Because of correlations of various characteristics, only item length was considered in the analysis of individual item characteristics.)

Table 1.4. Variables considered by Musseler (1994)

Category	Variable	Description/Prediction
Menu structure	position of item in menu	the farther to the right, the greater the selection time as predicted by Fitts' Law
	number of submenu items	decision time was predicted by Hick-Hyman law, movement time by Fitts' Law
Cognitive structure of submenu	fanning	action fanning (delete file, delete line, etc.) vs. object fanning (load file, delete file, etc.)
	overt vs. covert	overt (part of submenu name in menu name: load file under file) vs. covert (submenu name not in menu name: exit editor in file menu)
	grouping	present--submenu consists of print block, delete block, etc. and replace word, copy word, etc. or absent, more segmentation increases selection time
Individual items	item length	more characters take longer to read
	length of menu bar	more options take longer to read
	width of submenu	longest item in a submenu (more characters) takes longer to read

Table 1.5. shows the selection time equation coefficients as can be best determined by the authors. For example, the selection time (ms) for a menu bar while using a mouse is given by the equation:

time (ms) = 68 (position of item in menu) + 9 (# of submenu items)
+ fanning effect + overt/covert factor + 595 (# of subgroups)
+ 93 (item length).

In several cases, how variables should be coded is not explained, so what is shown represents the authors' best guess. For example, for position number in the menu, items could have been coded 1-6 or 0-5.

Table 1.5. Regression coefficients (ms)

Task	Using a mouse		Using letter keys	
Selection time for	menu bar	submenu	menu bar	submenu
position number of item in menu (probably 1-6 or 0-5)	68	12	78	7
position of item in submenu (1-7)	9	125	-18	-73
number of submenu items (3-7)	126	-25	-242	197
fanning - none	0	0	0	-0
action	123	33	-30	103
object	115	299	812	115
number of groupings in a submenu	595	-101	426	178
overt vs. covert	0	0	0	0
	426	-136	-1524	515
item length (number of characters)	93	15	99	21

Notice that the factors having the biggest impact on selection time were: (1) the number of groupings in the submenu and (2) whether the menu item was overt or covert. In the case of letter keys, several submenu characteristics affected menu bar item selection time. Subjects appeared to plan ahead, thinking about submenu choice while planning menu choices.

Thus, the lack of consistency across studies makes developing predictions for new applications quite difficult. All of the cited research used static computer menus. None required time-sharing as would be the case in a moving motor vehicle. It was for these reasons that new, automotive-specific data was required.

1.3 OVERVIEW OF EXPERIMENTS CONDUCTED

In light of the previously listed issues, three experiments were conducted to investigate the usability of hierarchical menu interface for secondary functions by drivers. During the initial project planning, there was concern that in-vehicle menu selection experiments might pose unacceptable risks to drivers and might be difficult to schedule because some experiments could occur during the winter months when driving conditions are poor in Michigan. Accordingly, an experiment was conducted to determine if data collected in the UMTRI driving simulator was comparable to data collected on the road. While Reed and Green (1995) report such a comparison for the UMTRI simulator, improvements had been made to the simulator since that study was conducted, so the expected performance differences were uncertain.

The second experiment examined the effects of learning on driving and performing a secondary task concurrently in the simulator. At issue was how much practice was required to stabilize primary task performance and if secondary task data from the early phases of practice could be used to predict data from later phases when those tasks were well learned.

The third experiment, much larger in scope, examined the usability of alternative input devices and menu structures for hierarchical menu systems. Evaluation criteria included minimizing the visual demands of secondary control systems, minimizing interference with the primary steering task, and satisfying driver desires.

2. EXPERIMENT 1 - VALIDATION OF A DRIVING SIMULATOR FOR SECONDARY TASK EVALUATION

2.1 BACKGROUND AND ISSUES

Although the use of driving simulators is increasingly popular in automobile human factors research (Green, 1995a), simulator validity should be established by comparing on-road performance with in-simulator performance. Driving simulators have several potential advantages over on-the-road experiments. Studies that are too hazardous for the road can be conducted safely in a simulator, and the greater repeatability of experimental conditions lead to more consistent and reliable results. Furthermore, the range of experimental variations possible in a simulator maybe greater than that on the road. However, driving simulators have several disadvantages. Due to the lack of risk, drivers may drive in a less realistic and more erratic manner. In addition, due to the inadequate motion and visual cues, drivers are more likely to suffer from motion sickness. Furthermore, the reduced field of view and short accommodation distance in simulators can limit the application of simulator research.

To determine when these disadvantages are critical, validation of driving simulators is necessary. Blaauw (1982) mentioned two aspects of simulator validity examined in the previous studies. The first concerned the physical correspondence between the contexts (road geometry, vehicle motion, etc.). The second aspect focused on the correspondence between driver behavior (driving task and secondary task performance) on the road and in the simulator. The latter issue was assumed to be more important for simulator validity in human factors studies.

Relatively few studies have dealt with the behavioral correspondence in simulators. Blaauw (1982) conducted a study comparing driving performance on the road and in the simulator. Both experienced and inexperienced subjects drove on straight roads in both contexts. The results suggested that the simulator offered good absolute (on-the-road vs. in-the-simulator) and relative (within individual contexts) validity for longitudinal vehicle control, but only good relative validity for lateral vehicle control. A lack of absolute validity for the lateral vehicle control was due to the diminished perception of lateral translations (absence of kinesthetic feedback in the fixed-base simulator).

Reed and Green (1995) examined the effects of an in-vehicle secondary task (dialing a cellular phone) on driving performance, while driving on straight roads in both contexts. The results of absolute and relative validity of driving performance were consistent with that obtained from previous studies. Lateral (lane keeping) and longitudinal (speed keeping) controls were less precise in both contexts while performing the phone task, and secondary task performance decrement was greater in the simulator.

Kurokawa and Wierwille (1990) compared the secondary task performance of 17 conventional automotive instrument panel (IP) tasks on the road and in the simulator. The results suggested that eye fixation measures (IP glance duration and number of glances) showed little significant difference (1 out of 17 secondary task categories) between driving on the road and in the simulator. Forward-viewing time was longer in the simulator than on the road, and task completion time was significantly longer (7 out of 17 secondary task categories) in the simulator than on the road.

The experiment described in this report examined the simulator validity for secondary task evaluation. Both driving and secondary task performance were compared between a real vehicle and a driving simulator. In light of the above studies, the following issues were addressed in this experiment:

(1) How should the UMTRI driving simulator be configured to be an appropriate context for secondary task evaluation?

(2) How does the secondary task affect driving performance in both contexts?

(3) What is the absolute and relative validity of the driving simulator?

2.2 TEST PLAN

In this experiment, subjects drove both the UMTRI Driver Interface Research Simulator (1995 version) and, on the road, an instrumented car (1992 Honda Accord station wagon). Subjects concurrently performed the same secondary in-vehicle secondary task while driving. Differences in driving and secondary task performance between contexts were examined.

2.2.1 Subjects

A total of eight licensed drivers participated in the context comparison (validation) experiment. Within each age group (young: ages 20 to 30; old: 65 and over) there were 4 experienced drivers, 2 men and 2 women. All subjects were familiar with the test route for the on-the-road portion of the experiment. None of the subjects were familiar with the UMTRI driving simulator.

2.2.2 Roads Driven

Four different roads were selected in the Ann Arbor, Michigan (USA) area for the on-the-road portion of this experiment (Figure 2.2.1). The total time to drive through the whole route was approximately one hour. Table 2.2.1 provides a description of roads driven.

Table 2.2.1. Road description for the on-the-road portion

Road type	Road description	Speed limit (mi/hr)	Authors' estimate of driving workload
Rural road	two-lane straight road with light and smoothly flowing traffic	40	low
Residential road	two-lane winding and hilly road with light traffic flow	30	moderate
Suburban street	four-lane straight road with heavy traffic, four traffic signals	45	high
Expressway	four-lane intersection with moderate but smoothly flowing traffic, two entrances and exits	55	low

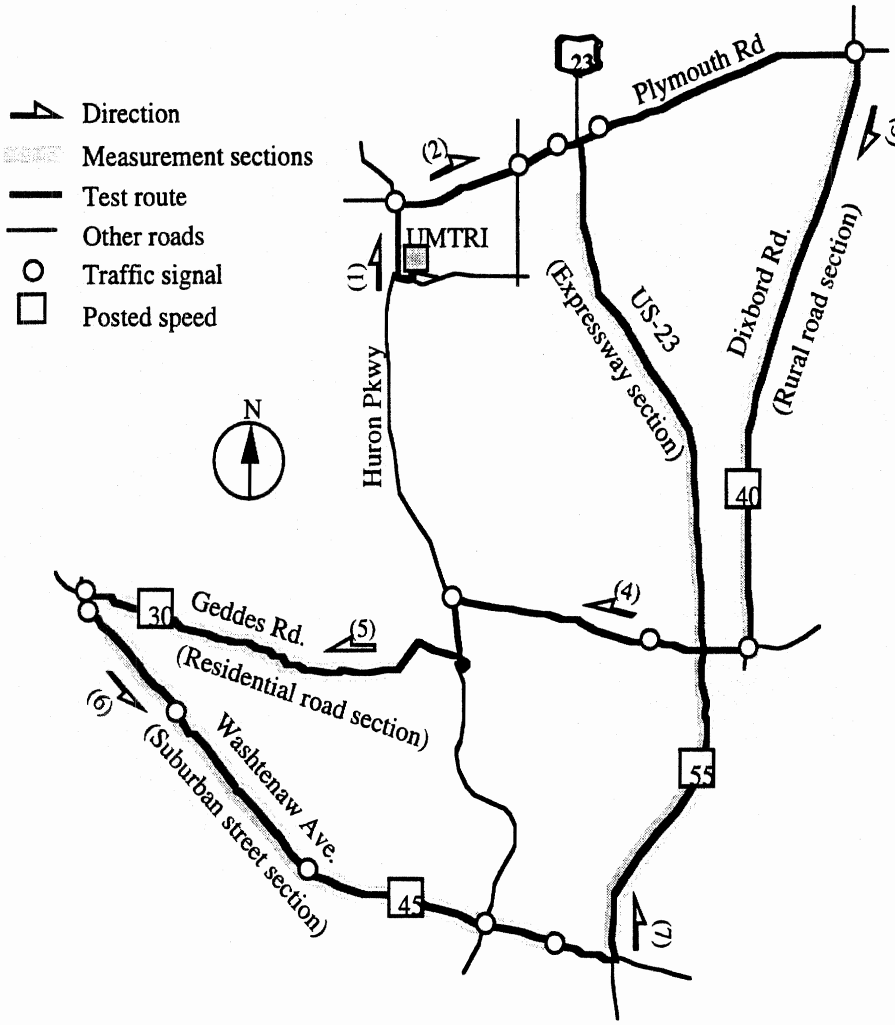


Figure 2.2.1. Test route for the on-the-road portion of the experiment

The same two-lane winding road (consisting of straight and curved sections) was used for all simulator conditions. (See Figure 2.2.2 for a typical road scene). The road was perfectly flat, and there were no lead vehicles, pedestrians, intersections, stop signs, or traffic lights. For reasons of cost and time, no effort was made to build a simulated road that matched the actual road driven in geometry. Instead, workload was manipulated by varying the speed driven in the simulator (35, 50, and 75 mi/hr) . The total time to drive in the simulator was approximately one hour. See the Appendix for more detailed information.

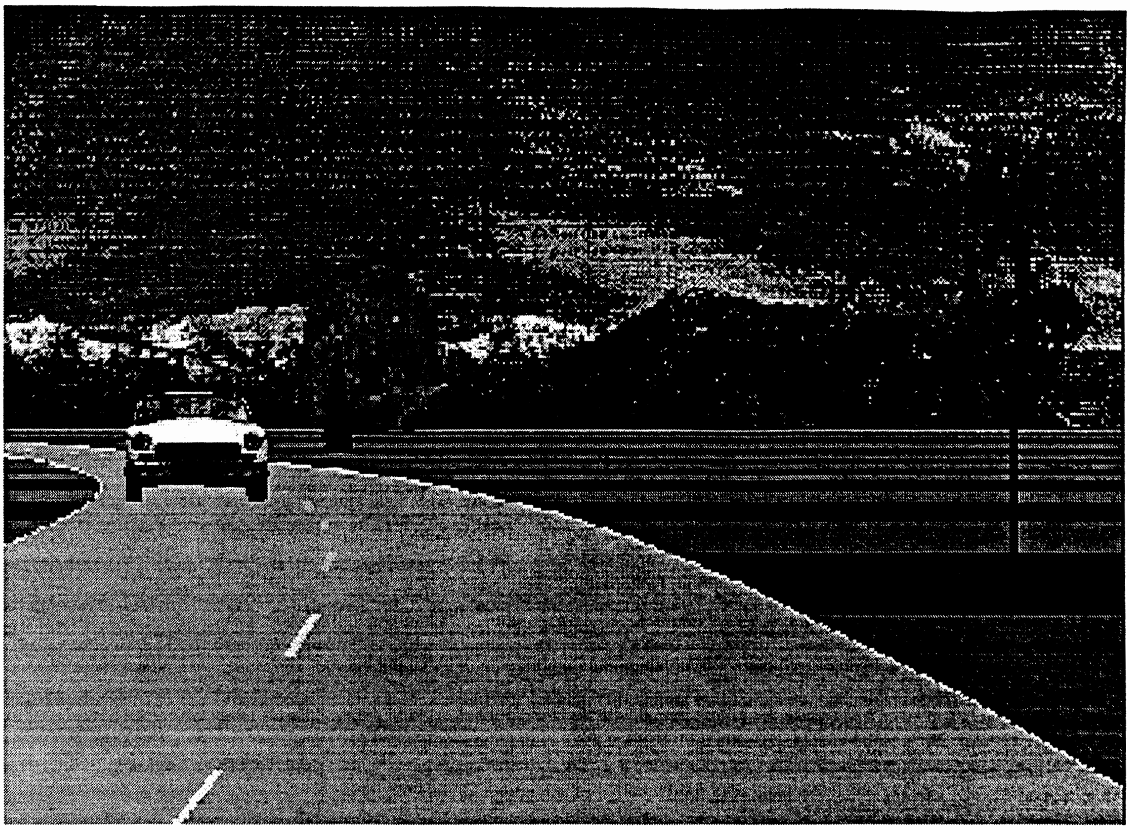


Figure 2.2.2. Typical simulated road scene

2.2.3 Instrumented Car

The test vehicle was a Honda Accord station wagon (Figure 2.2.4). Driving performance data (steering angle, lane position, throttle position, vehicle speed) were recorded by a computer in the vehicle at 30 Hz. Secondary task performance data (eye fixations, hand movements) were recorded on videotape using two CCD cameras. For a more complete description of the vehicle, see Sweet and Green (1993) and Katz, Green, and Fleming (1995).

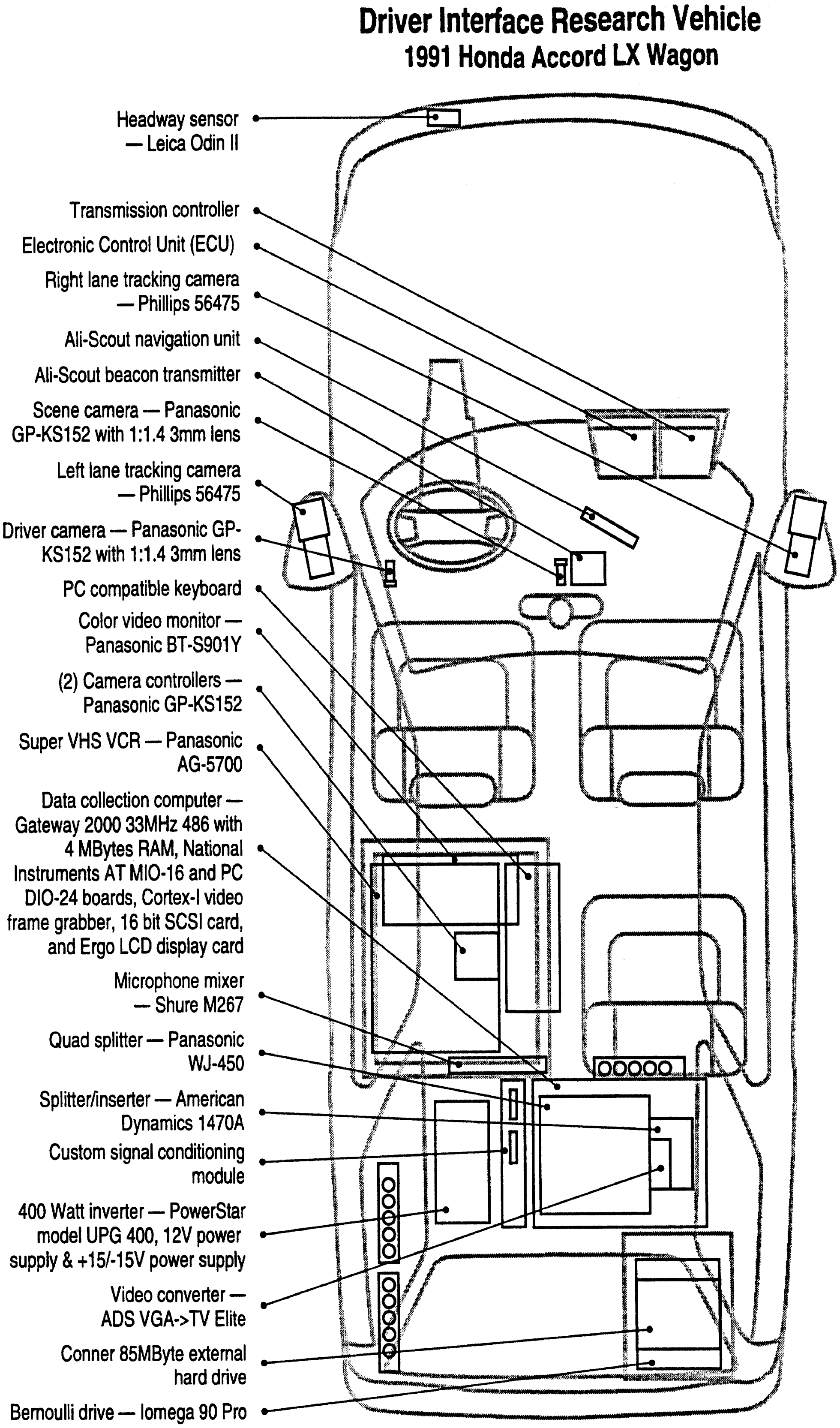
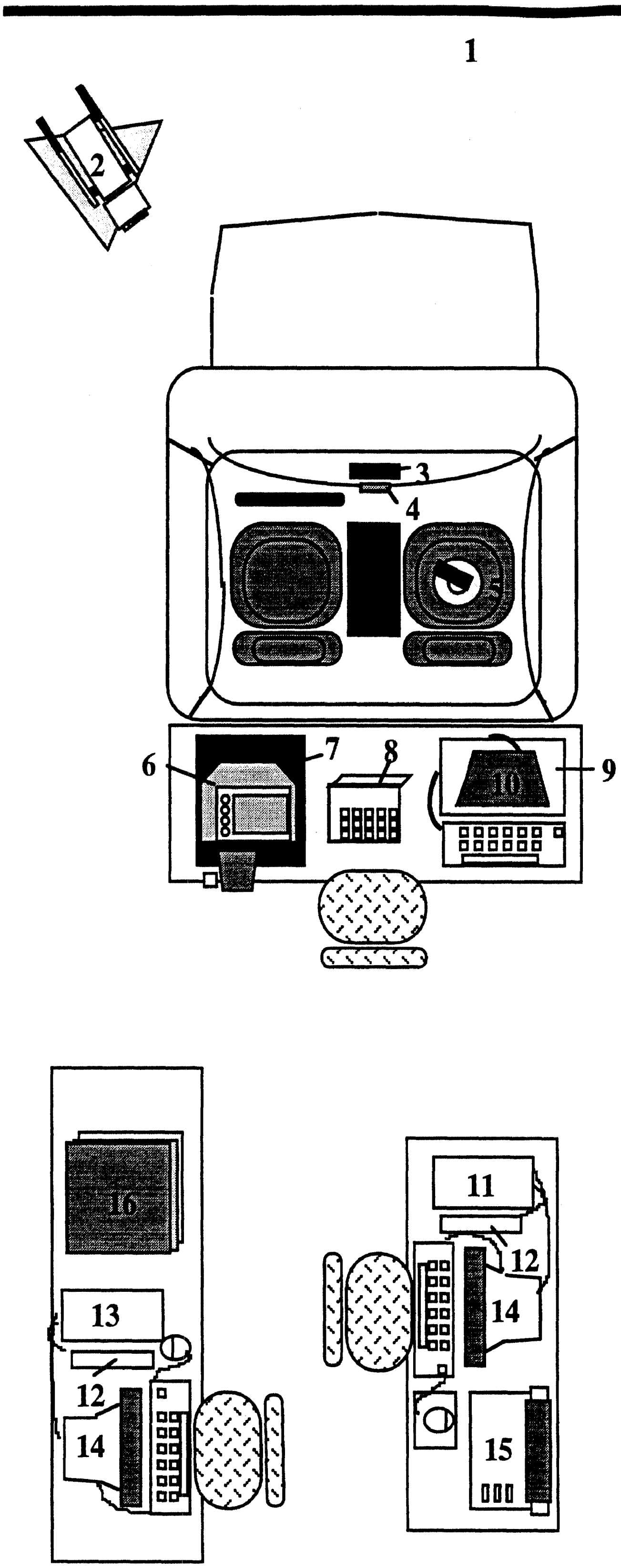


Figure 2.2.4. UMTRI instrumented vehicle

2.2.4 Driving Simulator

Laboratory data was collected using the UMTRI Driver Interface Research Simulator, a low-cost, fixed-base system based on personal computers. The simulator cab is from a 1985 Chrysler Laser and has an instrumented steering wheel, accelerator pedal, brake pedal, and a simulated speedometer (Figure 2.2.5). The computer-generated road scene was projected onto a screen (2.5 x 3.7 meters (8 x 12 feet)) located 6 meters (20 feet) in front of the cab. A sound system was used to simulate vehicle (engine) and driving noise (wind, off-road). Driving performance data was recorded by the main simulator computer. Secondary task performance data were recorded on videotapes using two CCD cameras. For a more complete description of the UMTRI driving simulator, see MacAdam, Green and Reed (1993) and Reed and Green (1995).

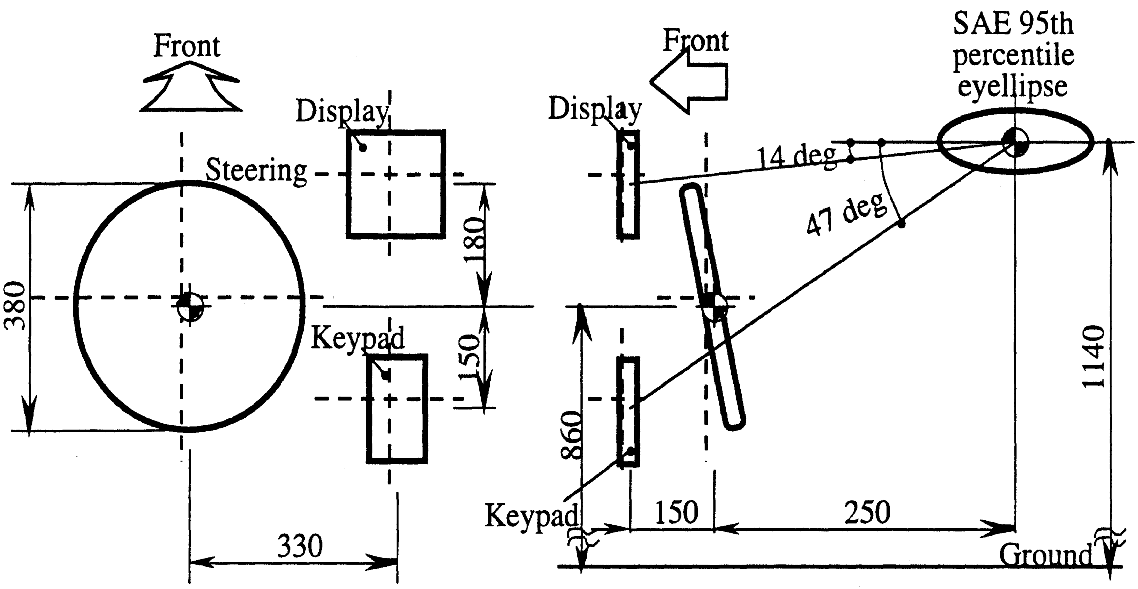


1 - Free-standing wall covered with 3M hi-white encapsulated reflective sheeting (8x12 ft)
2 - RCA model TC 1030/H10 low level light camera on tripod
3 - Hitachi model C5-LC2 5 inch LCD
4 - Kensington NoteBook keypad
5 - Panasonic model WJ-450 miniature camera
6 - Sharp model QA-1650 LCD computer projection panel
7 - 3M model 9000 RJH overhead projector
8 - Macintosh PowerBook Duo 270C
9 - IBM model 5160 personal computer with Orchid PC Turbo accelerator and Keytronic model KB 5151 keyboard
10 - IBM model 5151 Personal Computer display
11 - Power Macintosh 8100/80AV computer with Apple extended keyboard II and Apple Desktop Bus Mouse II
12 - Bernoulli model B190TM 90MB Mac transportable drive
13 - Macintosh Quadra 840AV computer with Apple extended keyboard II and Apple Desktop Bus Mouse II
14 - Macintosh model M1212 13" color display
15 - Apple model A9M0320 Imagewriter II
16 - Audio system
(from top to bottom)
Kenwood model KX-48C stereo cassette deck
Kenwood model KR-A4060 AM-FM stereo receiver
Kenwood model GE-7030 stereo graphic equalizer

Figure 2.2.5. Simulator setup for this experiment

2.2.5 In-Vehicle Secondary Task

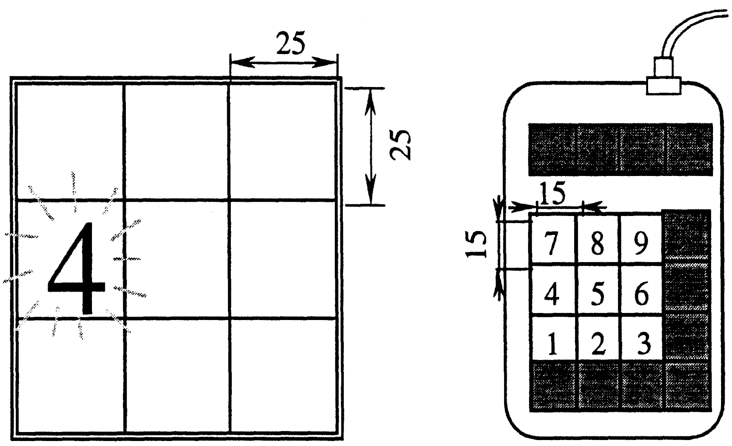
A number-selection task was used to simulate in-vehicle secondary systems. The same secondary task was employed both on the road and in the simulator. The equipment used consisted of a Hitachi 5-inch Liquid Crystal Display (LCD), a Kensington Note Book number keypad, and a Macintosh Duo 270C portable computer (which ran the secondary task program). The display and the number keypad were located at the top and bottom of the center console respectively (Figure 2.2.6).



(All dimensions displayed are in millimeters)

Figure 2.2.6. Locations of the LCD and the number display

A HyperCard program controlled the numbers presented on the display. The display was divided into a 3 x 3 matrix of squares. The number arrangement on the display corresponded to the configuration of the keypad (Figure 2.2.7). Secondary task data (interbutton response time and task completion time) were recorded by the portable computer.



(All dimensions displayed are in millimeters)

Figure 2.2.7. Sample display generated by HyperCard program and keypad configuration

2.2.6 Procedure Overview

The experiment was conducted in two parts, on the road and in the simulator. The order was counterbalanced across subjects. Both parts contained four block sequences, each sequence lasting five minutes: one for each different road type for the on-the-road part, and one for each different speed condition for the simulator, plus a practice block prior to actual data collection. Within parts, the order for road types and speed conditions was fixed for all subjects. Subjects were instructed to drive in a normal and safe manner, to maintain speeds at or below the posted speed limits on the road, and to drive at steady speed for each of the speed conditions in the simulator.

In both contexts, each subject completed a secondary task of entering five 5-digit sequences while driving in each road and speed condition (Figure 2.2.8).

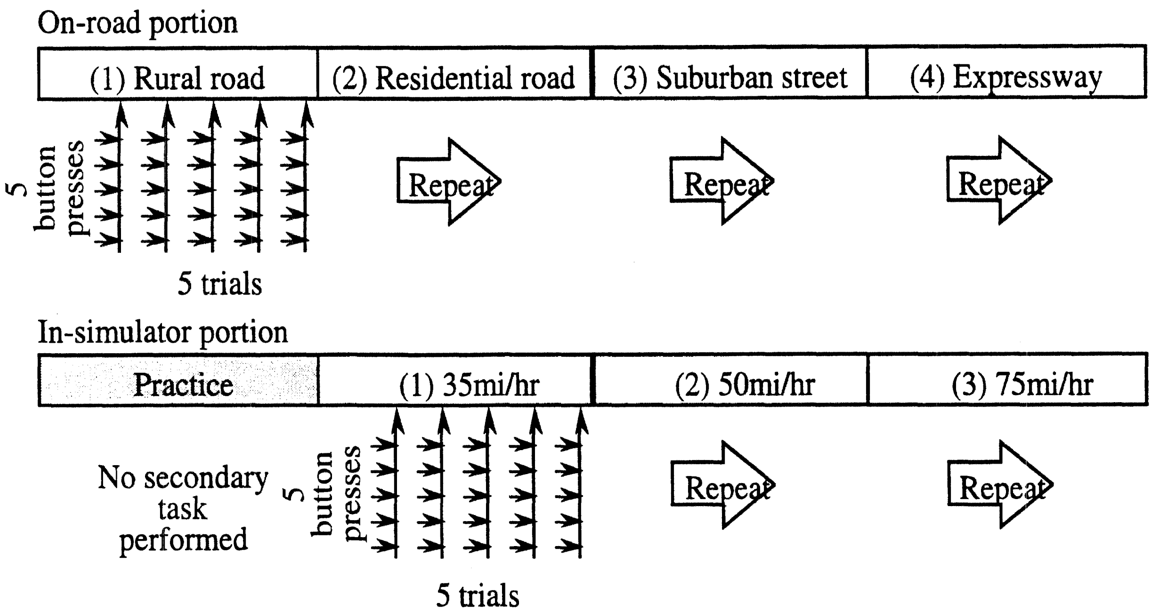


Figure 2.2.8. Procedure of the context comparison experiment

When the experimenter triggered the secondary task, a random number between 1 to 9 appeared in 1 of the 9 spatially compatible squares on the display, accompanied by a beep. Subjects pressed the button on the keypad corresponding to the number displayed. Immediately, another number appeared on the display and the subjects responded (Figure 2.2.9).

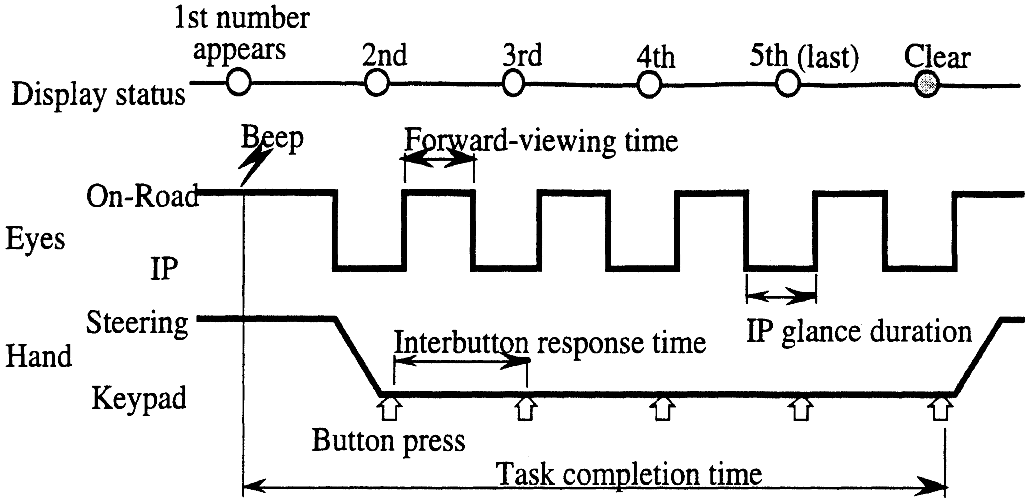


Figure 2.2.9. Secondary task timing and driver behavior

2.2.7 Experimental Design

Independent variables examined in this experiment (Table 2.2.2) were subjects nested within age and gender as well as context and secondary task (both between-subject effects). The four road types on the road and three speed conditions in the simulator were combined into a single independent variable (context), to facilitate comparing the on-road environment with the simulator. For the secondary task, all driving performance data was divided into two components: dual-task performance (driving with the secondary task), or single-task performance (driving without the secondary task).

Table 2.2.2. A summary of independent variables

Independent variable	Levels
Age (A)	young and old
Gender (G)	male and female
Subject (S)-nested	8 participants
Context (C)	on-the-road (4): rural, residential, suburban, expressway simulator (3): 35, 50, 75 mi/hr
Secondary task (T)	present and absent

As can be seen in Table 2.2.3, dependent measures were categorized into two groups: secondary task and driving performance. Eye fixation data was obtained from videotape and analyzed frame by frame (33 ms per frame). The driving performance variables were intended to measure input from the drivers (effort) and output from the vehicle (quality) to complete a closed-loop system.

Table 2.2.3. A summary of dependent measures

Task category	Data source	Dependent measure	Explanation
Secondary task performance	decision and movement time	interbutton response time (s)	time between button presses
		task completion time (s)	total time from first beep of sequence to completion of final button press
	eye fixation	IP glance duration (s)	both to the control and to the display
		number of IP glances	in a trial
		forward viewing time (s)	between successive glances to IP
Driving performance	lateral control	s.d. of steering angle (deg)	
		s.d. of standardized lane position (%)	percentages of vehicle position from road center to road edge
	longitudinal control	s.d. of throttle position (%)	
		s.d. of vehicle speed (mi/hr)	

2.3 RESULTS

An analysis of variance (ANOVA) test was used to investigate the effects of independent variables on secondary task and driving performance. Prior to the ANOVA tests, the histograms and various summary statistics (e.g., skew, kurtosis) of the acquired data were compared with those of the log (1+x) transformed data to verify the normality of distributions. The distributions of both the secondary task (excluding number of IP glances) and driving performance tended to be lognormal (Figures 2.3.1 and Figure 2.3.2), but the departure from normality was such that use of the untransformed data in the ANOVA model was appropriate.

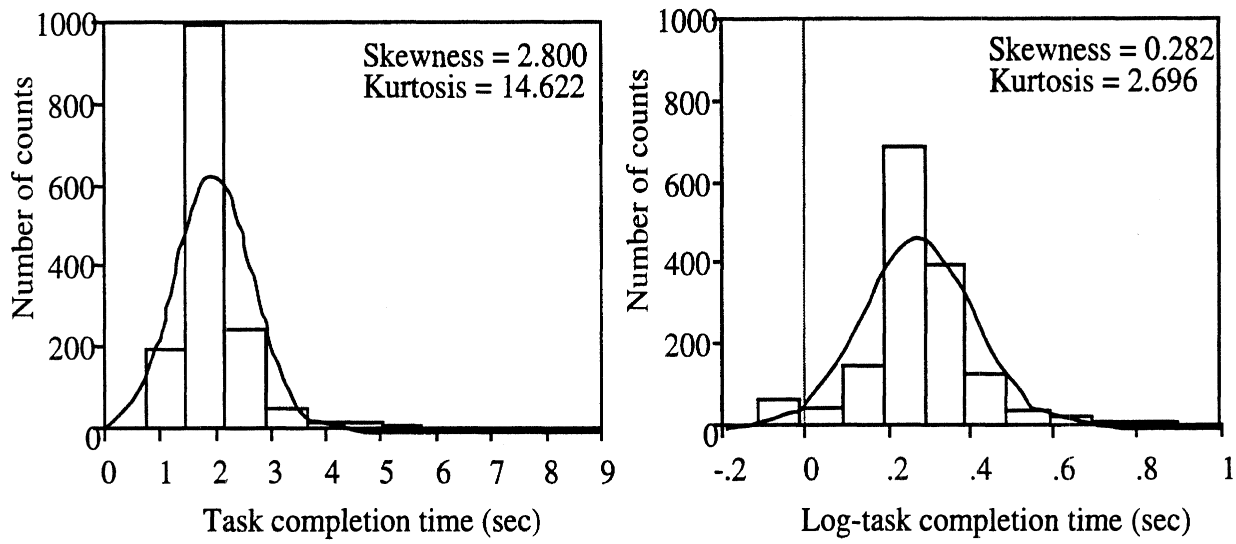


Figure 2.3.1. Histograms of task completion time

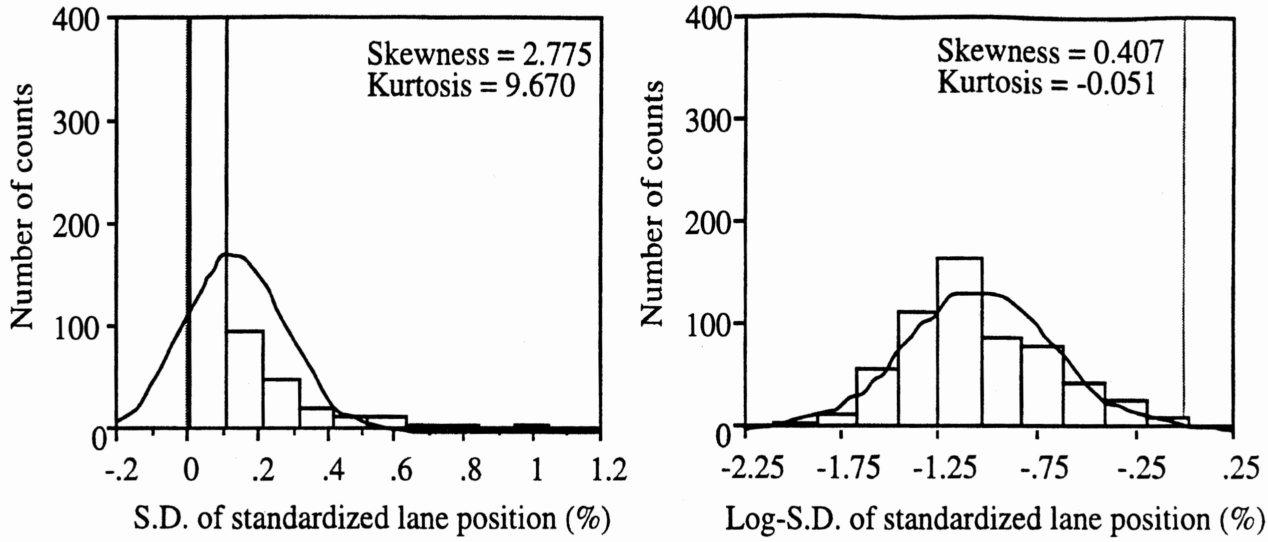


Figure 2.3.2. Histograms of the S.D. of standardized lane position

The data were analyzed in two steps. First, the simulator and on-road data were pooled, and all main effects and one-way interactions were examined in a single ANOVA. For both contexts, within-context effects (three speed conditions for the simulator, four road types on the road) were ignored. In the second step, separate fixed-effect ANOVAs were computed for the simulator data and the on-road data. This approach was chosen for computation ease, allowing the within-context effects and their interactions to be readily computed.

2.3.1 Secondary Task Performance (Combined Comparison)

From the ANOVA summary of combined comparison (Table 2.3.1), there were no significant differences in secondary task times between the two contexts, with on-road and simulator interbutton response times (RTs) and task completion times differing by one percent. The results indicated significant effects of age ($p < 0.01$), gender ($p < 0.01$) and significant age by gender interactions ($p < 0.01$) on hand movement measures. These results indicated that the older age group required about 25 percent longer to respond and to complete the same secondary task than the younger age group (2209 ms vs. 1780 ms for the interbutton RT, 11.503 sec versus 8.961 sec for task completion times). Individual subject differences were also statistically significant ($p < 0.01$).

Table 2.3.1. ANOVA summary of combined comparison

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++	++	++	++	++
Gender (G)	++	++	++	++	++			++	++
Subject (S)	++	++	++	++	++			++	++
Context (C)			++	++		++	++	++	++
Task (T)	n/a	n/a	n/a	n/a	n/a	++	++	++	++
A x G	++	++	++	+				+	++
A x C			++	+		+	++	++	++
A x T	n/a	n/a	n/a	n/a	n/a	++			
G x C								+	++
G x T	n/a	n/a	n/a	n/a	n/a				+
S x C	++		++	++	+		+	++	++
S x T	n/a	n/a	n/a	n/a	n/a				
C x T	n/a	n/a	n/a	n/a	n/a		++	++	

(++: $p < 0.01$; +: $p < 0.05$)

In terms of the glance-related measures, the results indicated significant effects of context ($p<0.01$) on IP glance duration and forward-viewing time (Figure 2.3.3). Subjects paid less attention to the IP and greater attention to the forward scene in the simulator than on the road. This occurred because no effort was made to match geometry of the simulated road (with a moderate number of curves) to read roads. No significant effect of context was found on the number of IP glances.

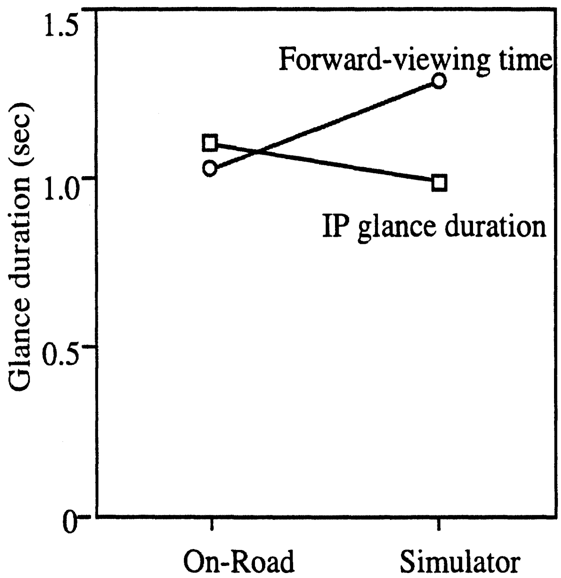


Figure 2.3.3. Effect of context on IP glance duration and forward-viewing time

The results indicated significant effects of age ($p< 0.01$) and gender ($p< 0.01$) on eye fixation measures. There was a significant age by gender interaction ($p< 0.01$) for IP glance duration and a marginally significant age by gender interaction ($p< 0.05$) for forward-viewing time. These results illustrated that the older age group needed to pay less attention to the IP and greater attention to the forward scene than the young age group.

The results indicated significant age by context interactions ($p < 0.01$) on IP glance duration (Figure 2.3.4), and a marginally significant age by context interaction ($p < 0.05$) on forward-viewing time (Figure 2.3.5). For the IP glance duration, the differences between age groups were smaller in the simulator than on the road. For the forward-viewing time, the differences between age groups were larger in the simulator. These results indicated that IP glance duration and forward viewing time had an inverse relationship between two contexts.

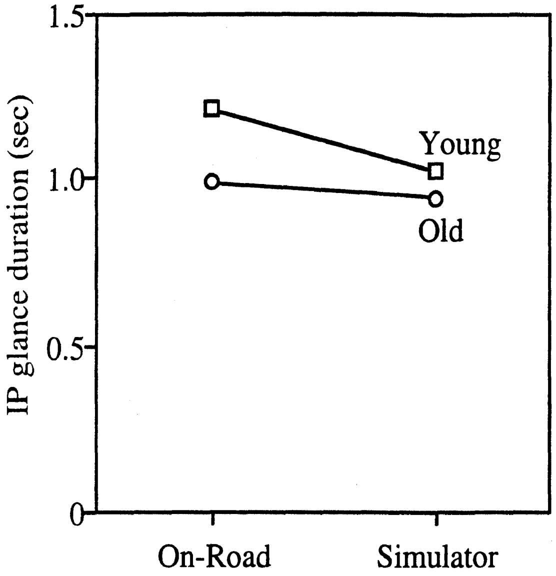


Figure 2.3.4. Age by context interaction of IP glance duration

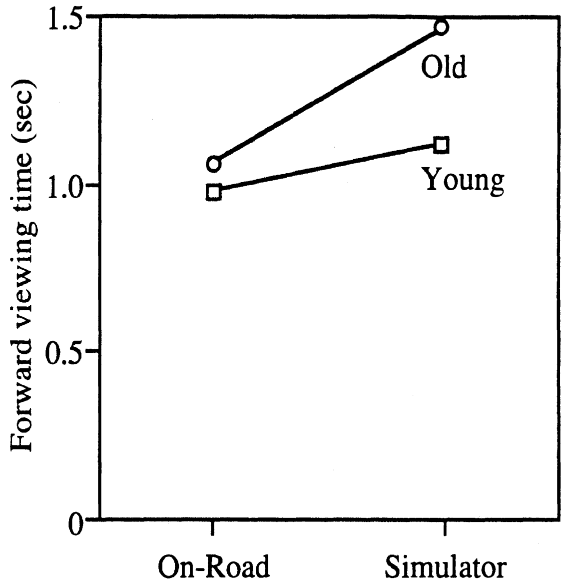


Figure 2.3.5. Age by context interaction of forward-viewing time

To provide a sense of the nature of the differences found, the relative differences between the condition means ((simulator-road) / road)) for each secondary task measures are shown in Figure 2.3.6. Positive values represented measures that were larger in the simulator than on the road. On the other hand, negative values represented situations where values were larger for the simulator. The results illustrated that hand movement measures (interbutton RT and task completion time) and the number of IP glances were unaffected by the differences between the contexts. However, eye fixation measures (IP glance duration and forward-viewing time) were influenced by the contexts and showed an inverse relationship.

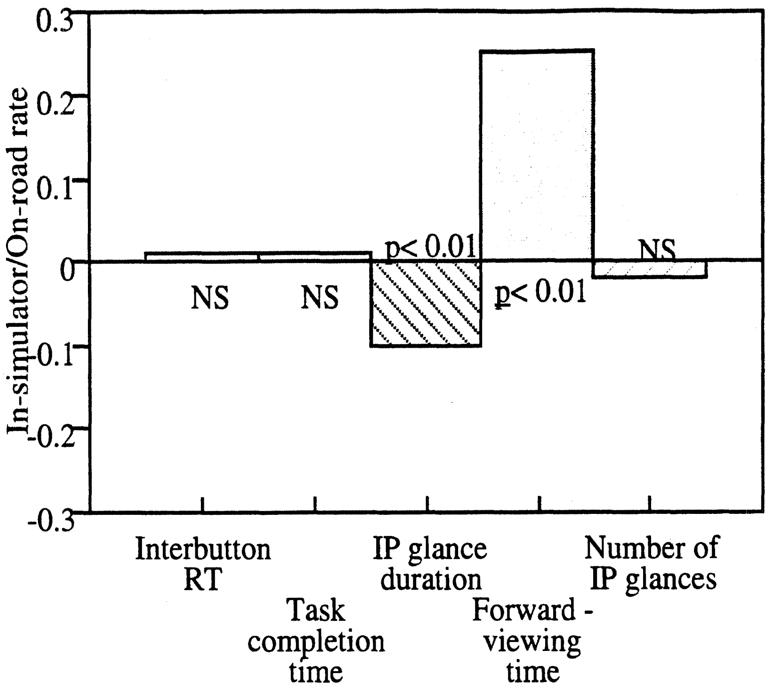


Figure 2.3.6. Summary effects of context on secondary task performance

2.3.2 Driving Performance (Combined Comparison)

As with the other measures examined (Table 2.3.1), there were significant differences due to age ($p < 0.01$) for the two lateral control measures (S.D. of steering wheel angle, S.D. of lane position) and the longitudinal control measures (S.D. of throttle angle, S.D. of vehicle speed), but there were gender and subject differences only for longitudinal control. The combined comparison also indicated significant effects of context ($p < 0.01$) on the S.D. of steering angle (21.3 in the simulator versus 3.4 on the road), the S.D. of standardized lane position (simulator=20.9, road=5.9), the S.D. of throttle position (simulator=19.6, road=6.9), and the S.D. of vehicle speed (simulator=2.7, road=3.4). In brief, steering wheel angle variability and lane position variability were both greater on the road (as one would expect) because the simulated road had more curves. In contrast, throttle variance was greater on the road but speed variance was less. Figure 2.3.7 shows the normalized differences. As before, positive values correspond to larger average values in the simulator. Again, readers should bear in mind that the conditions were not matched. In fact, the simulator data are for three different driving conditions and the on-road data are for four, very different roads.

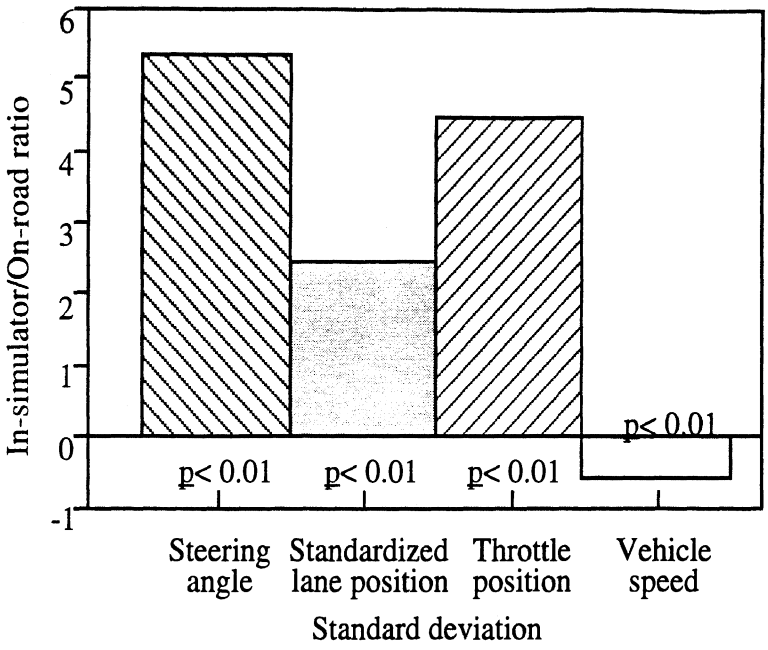


Figure 2.3.7. Summary effects of context on driving performance

There was a marginally significant age by context interaction ($p < 0.05$) for the S.D. of steering angle (Figure 2.3.8), and there were significant age by context interactions ($p < 0.01$) for the S.D. of standardized lane position (Figure 2.3.8), the S.D. of throttle position (Figure 2.3.10), and the S.D. of vehicle speed (Figure 2.3.11). Differences between age groups for the lateral and longitudinal control measures were considerable in the simulator. The older age group showed greater variance than the young age group. However, age group differences were minimal on the road. This age-related interaction may reflect the ease with which each age group can learn a task with new elements (in this case, a simulator).

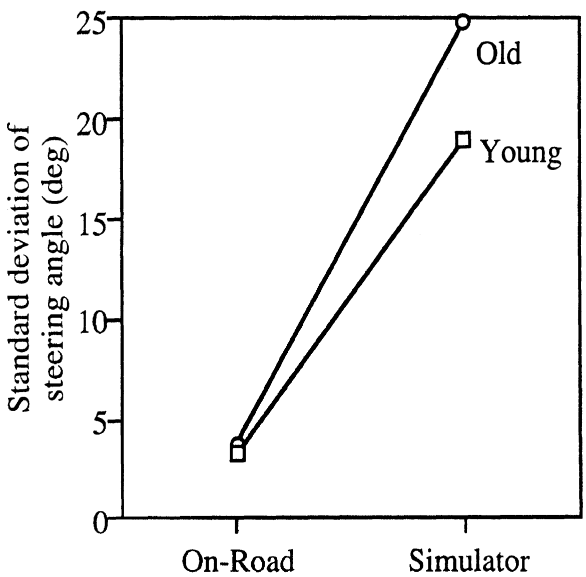


Figure 2.3.8. Age by context interaction on S.D. of steering angle

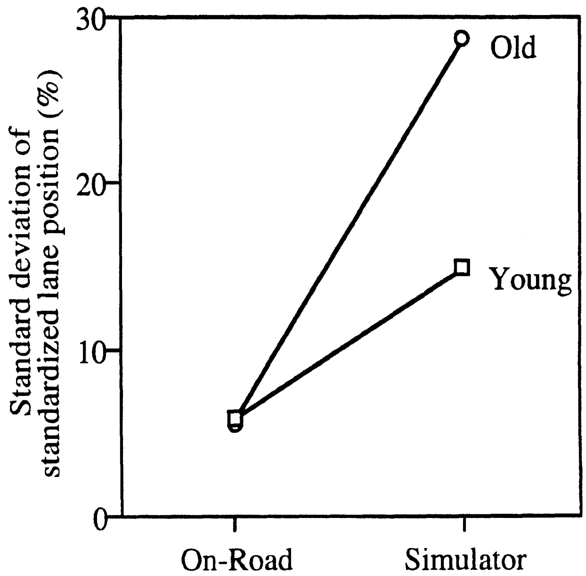
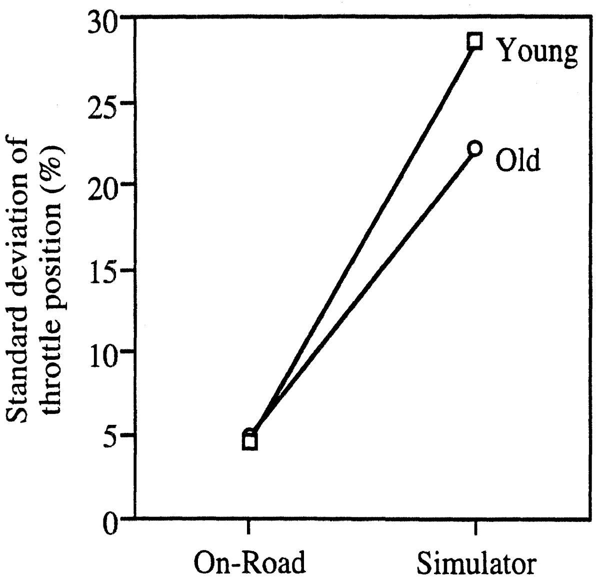


Figure 2.3.9. Age by context interaction on S.D. of standardized lane position



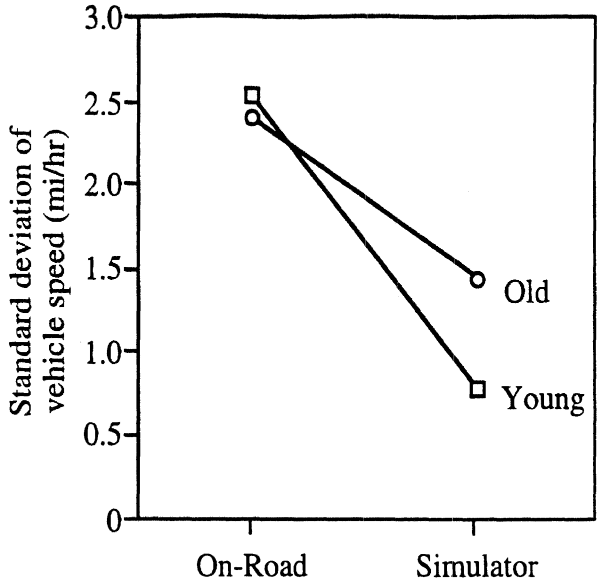
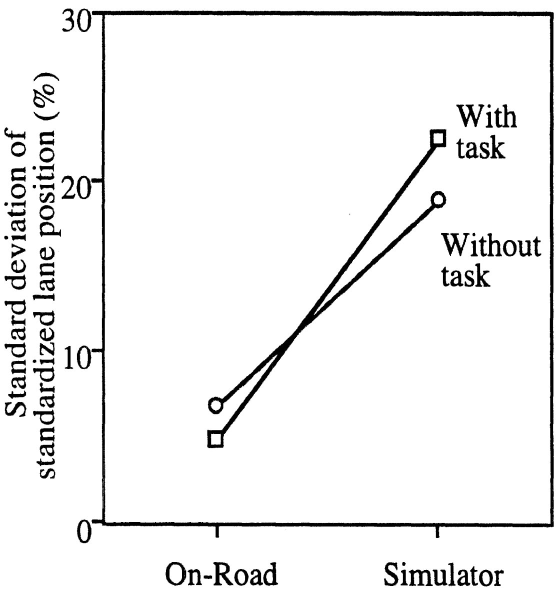


Figure 2.3.10. Age by context interaction of S.D. of throttle position

Figure 2.3.11. Age by context interaction of S.D. of vehicle speed

For two of the driving performance measures, S.D. of standardized lane position (Figure 2.3.12) and the S.D. of throttle position (Figure 2.3.13), there were significant context by secondary task interactions ($p < 0.01$) with the simulator appearing to accentuate the absolute magnitude of differences between task and no task conditions. These results suggested that the S.D. of standardized lane position in the simulator could be a sensitive measure for evaluating the secondary task impacts.



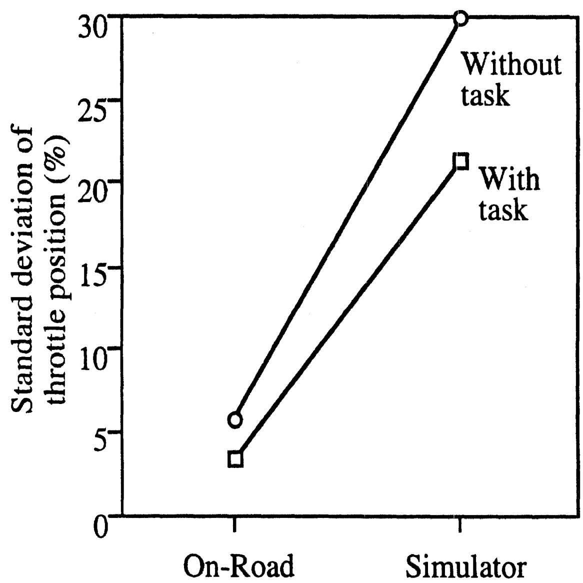


Figure 2.3.12. Context by secondary task interaction of S.D. of standardized lane position

Figure 2.3.13. Context by secondary task interaction of S.D. of throttle position

2.3.3 Secondary Task Performance (Separate Analyses)

From the separate ANOVAs of the road (Table 2.3.2) and in the simulator (Table 2.3.3) data, the results indicated significant effects of context ($p < 0.01$) on interbutton RT (Figure 2.3.14) and task completion time (Figure 2.3.15) on the road. In contrast to the combined analyses described earlier, road context effects refer to differences between roads (as opposed to the simulator versus the road in earlier analyses). There were no significant effects of context (between speed conditions) in the simulator. A Scheffe's test on interbutton RT showed significant differences in road types ($p < 0.01$) between rural road and suburban street, and between expressway and suburban street. A Scheffe's test on task completion time showed a marginally significant difference in road types ($p < 0.05$) between rural road and suburban street. On the suburban street, subjects needed less time to respond and to complete the secondary task.

Table 2.3.2. ANOVA summary of individual comparison (on-road)

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++				
Gender (G)			++	++	++				
Subject (S)	++	++	++	++	++		+	++	
Context (C)	++	++	+			++		+	++
Task (T)	n/a	n/a	n/a	n/a	n/a	++	++	++	++
A x G	++	++	++	++				++	+
A x C	++	+			+				
A x T	n/a	n/a	n/a	n/a	n/a				
G x C			+	+				+	++
G x T	n/a	n/a	n/a	n/a	n/a				+
S x C	++		++	++					
S x T	n/a	n/a	n/a	n/a	n/a				
C x T	n/a	n/a	n/a	n/a	n/a				+

(++: $p < 0.01$; +: $p < 0.05$)

Table 2.3.3. ANOVA summary of individual comparison (simulator)

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++	++	++	++	++
Gender (G)	+	+	++	+	++			++	++
Subject (S)	++	++	++	++	++			++	++
Context (C)			++	+		++	++	+	++
Task (T)	n/a	n/a	n/a	n/a	n/a	++			++
A x G	+		++					++	++
A x C									
A x T	n/a	n/a	n/a	n/a	n/a	++			
G x C									++
G x T	n/a	n/a	n/a	n/a	n/a			+	+
S x C	+							++	++
S x T	n/a	n/a	n/a	n/a	n/a		+		
C x T	n/a	n/a	n/a	n/a	n/a				

(++: p< 0.01; +: p< 0.05)

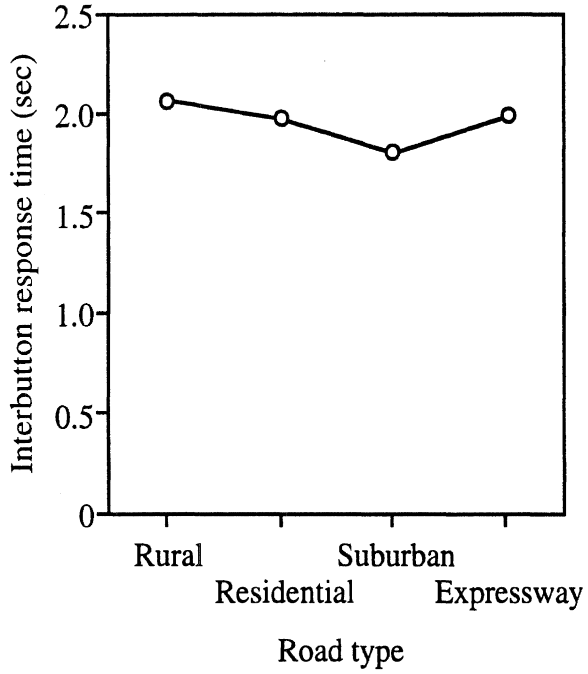


Figure 2.3.14. Effect of context on interbutton response time (on the road)

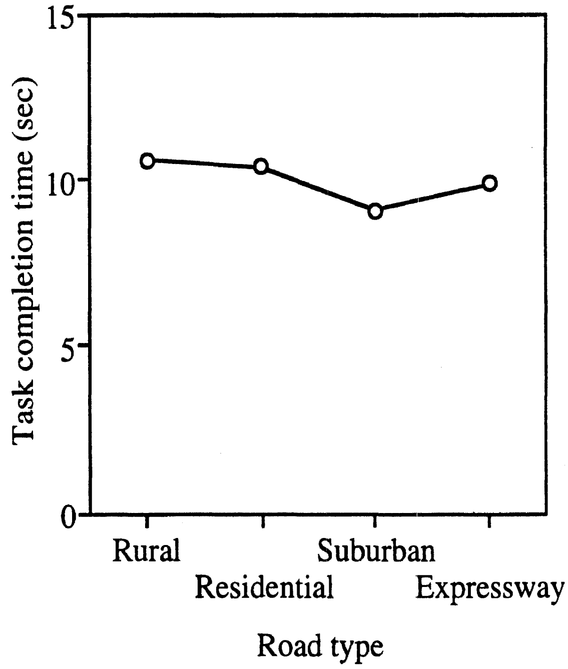


Figure 2.3.15. Effect of context on task completion time (on the road)

The results on the road showed a marginally significant effect of context ($p < 0.05$) on IP glance duration and no significant effect of context on forward-viewing time (Figure 2.3.16). The results in the simulator showed a significant effect of context ($p < 0.01$) and a marginally significant effect of context ($p < 0.05$) on forward viewing time (Figure 2.3.17). A Scheffe's test on the road showed no significant difference between road types at this level on IP glance duration. A Scheffe's test in the simulator showed a marginally significant ($p < 0.05$) difference between 35 mph and 75 mph on IP glance duration, and a marginally significant difference ($p < 0.05$) between 35 mph and 50 mph, and between 35 mph and 75 mph on forward-viewing time. An inverse relationship between IP glance duration and forward-viewing time in the simulator was found as a function of the three speed conditions in the simulator. There were no significant effects of context on number of IP glances in both contexts.

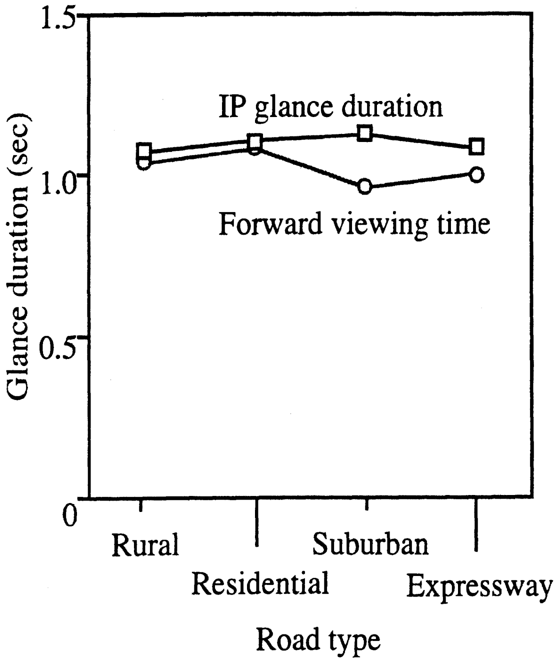


Figure 2.3.16. Effect of context of IP glance duration and forward-viewing time on the road

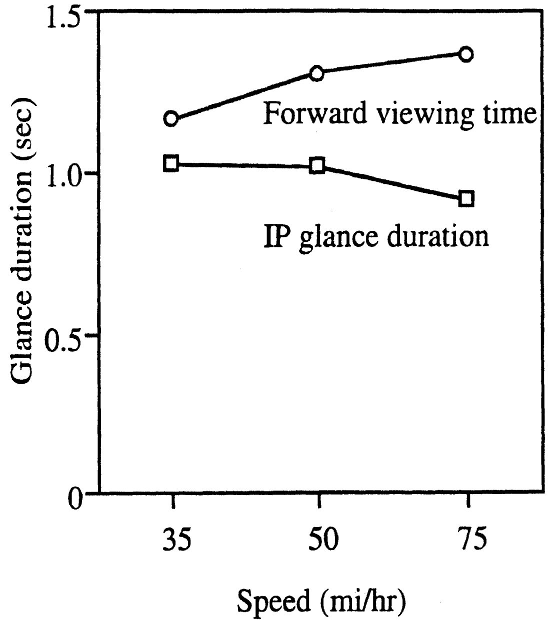


Figure 2.3.17. Effect of context of IP glance duration and forward-viewing time in the simulator

2.3.4 Driving Performance (Separate Analyses)

From the ANOVA summary of individual comparison on the road (Table 2.3.2) and in the simulator (Table 2.3.3), the results for the S.D. of steering angle indicated a significant effect of context ($p < 0.01$) on the road (Figure 2.3.18) and in the simulator (Figure 2.3.19) . A Scheffe's test on the road indicated significant differences in road types ($p < 0.01$) between residential road and others road (rural, suburban, expressway). On the winding residential road, larger steering wheel angles were sometimes needed. A Scheffe's test in the simulator showed significant differences in speed conditions ($p < 0.01$) among all combinations. These results illustrated a linear relationship between steering wheel angle variability and speed in the simulator.

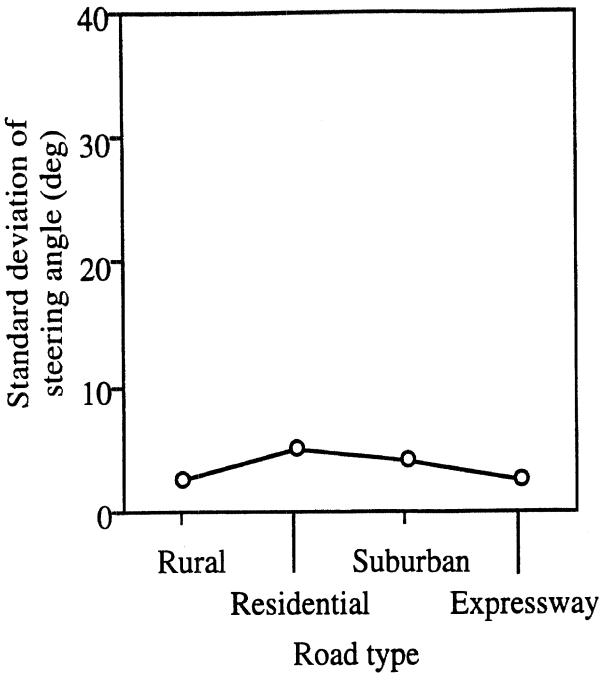


Figure 2.3.18. Effect of actual road driven on S.D. steering angle

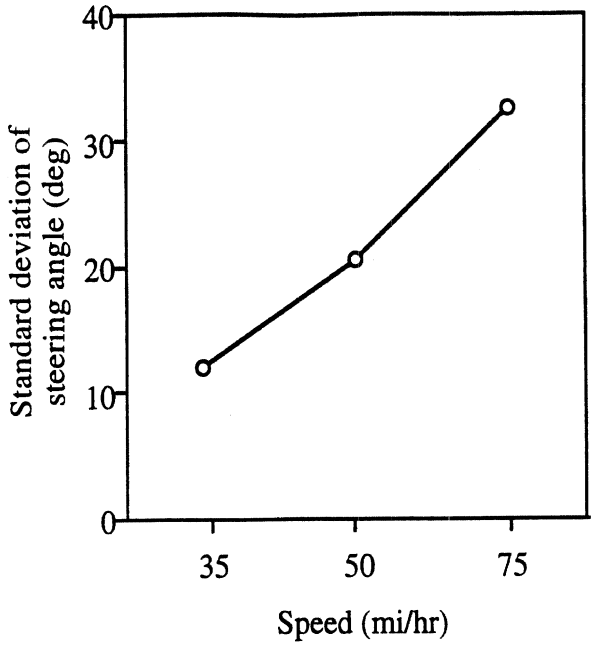


Figure 2.3.19. Effect of speed in simulator on S.D. of steering angle

The results for the S.D. of standardized lane position indicated no significant effect of context on the road (Figure 2.3.24) and a significant effect of context ($p < 0.01$) in the simulator (Figure 2.3.25). A Scheffe's test in the simulator showed significant differences in speed conditions ($p<0.01$) among all combinations. The same linear relationship observed in the S.D. of steering angle was shown in drivers' lane-tracking performance in the simulator.

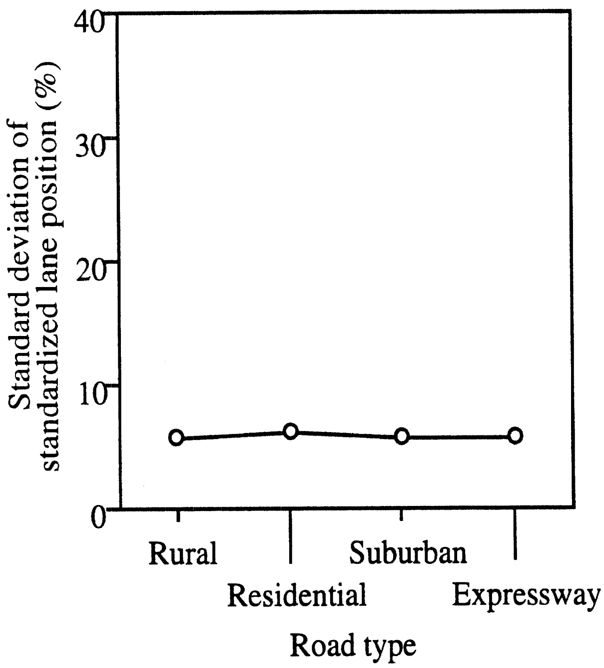


Figure 2.3.20. Effect of actual road driven on S.D. of standardized lane position

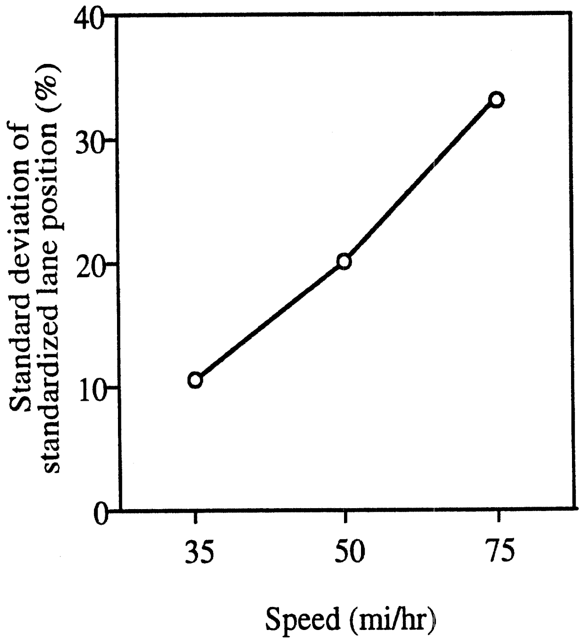


Figure 2.3.21. Effect of speed in simulator on S.D. of standardized lane position

The results for the S.D. of throttle position indicated marginally significant effects of context ($p<0.05$) on the road (Figure 2.3.22) and in the simulator (Figure 2.3.23). A Scheffe's test on the road showed a significant difference in road types ($p< 0.05$) between residential road and suburban street. A Scheffe's test in the simulator showed a significant difference in speed conditions ($p<0.01$) between 35 mph and 75 mph.

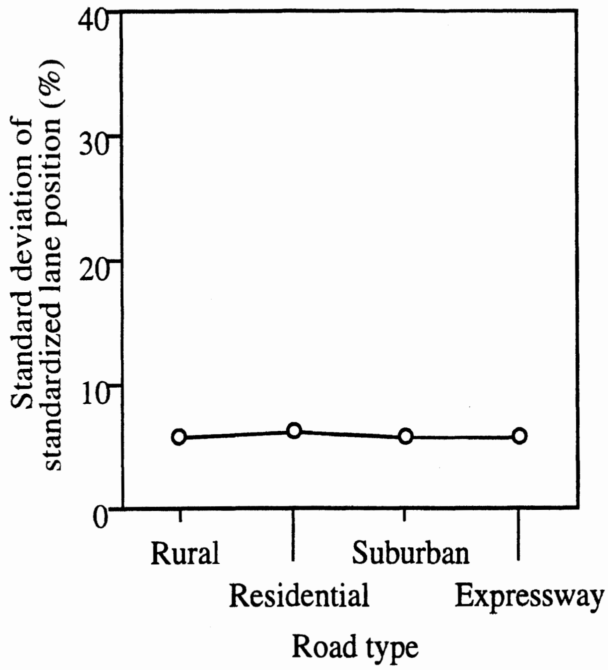


Figure 2.3.22. Effect of actual road driven on S.D. of throttle position

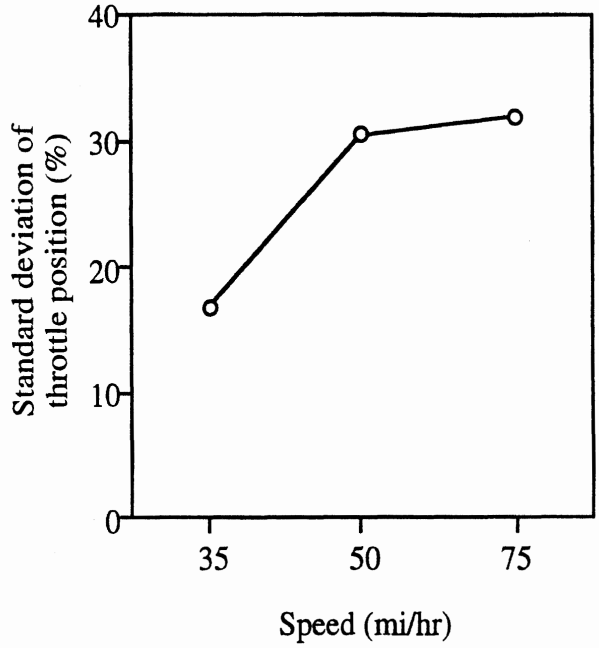


Figure 2.3.23. Effect of speed in simulator on S.D. of throttle position

The results for the S.D. of vehicle speed indicated a significant ($p< 0.01$) differences between road types on the road (Figure 2.3.24) and speeds in the simulator (Figure 2.3.25). A Scheffe's test of the on-the-road data indicated significant ($p< 0.01$) differences between suburban street (having the greatest speed variability) and other roads (rural, residential, expressway), and between rural road and expressway ($p< 0.05$). A Scheffe's test of the simulator data showed significant differences ($p< 0.01$) between all combinations of speed conditions.

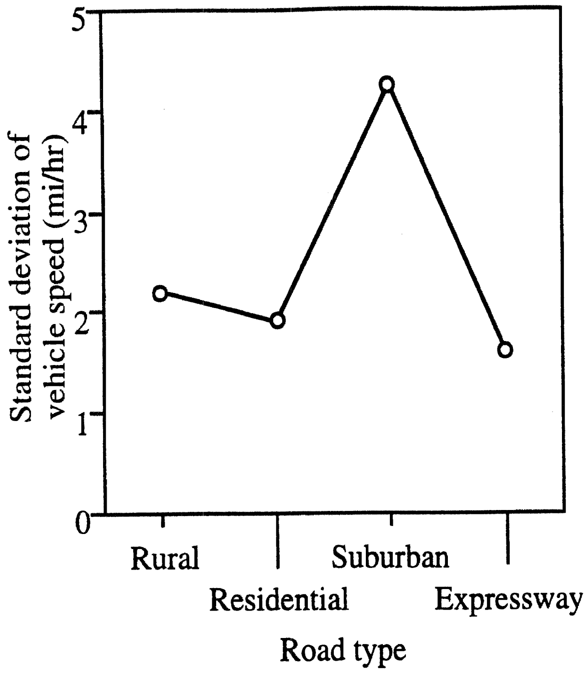


Figure 2.3.24. Effect of actual road driven on S.D. of vehicle speed

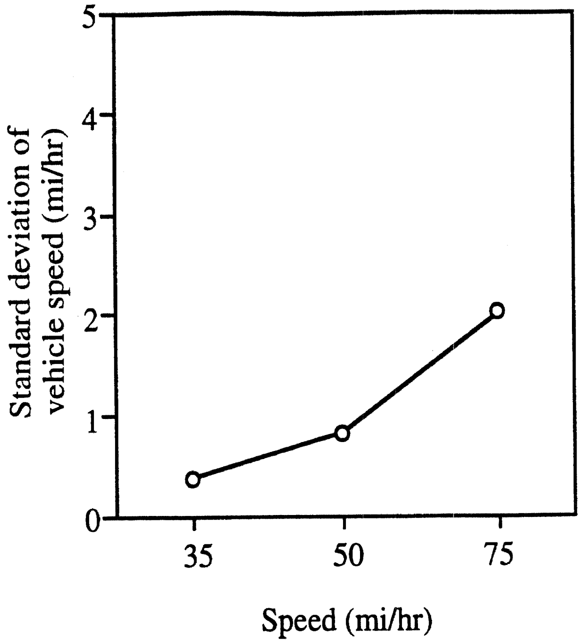


Figure 2.3.25. Effect of speed in simulator on S.D. of vehicle speed

The ratios of the driving performance measures with and without secondary task are summarized (Table 2.3.34). Positive values represented that the performance was worse (larger standard deviation) with a secondary task than without a secondary task. The results indicated that only the S.D. of standardized lane position in the simulator had a larger variance with a secondary task than without a secondary task.

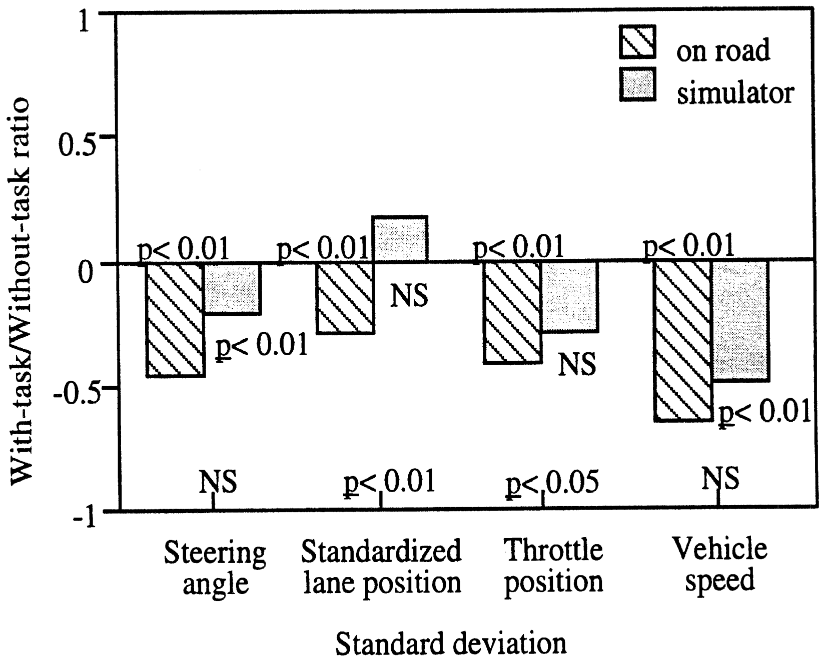


Figure 2.3.26. Summary effects of secondary task on driving performance

2.4 DISCUSSION

The results of secondary task performance showed no difference in task time measures (interbutton RT, task completion time) between the simulator and on-the-road driving. In both contexts, the time needed for the drivers to respond and to complete the secondary task was the same. This finding indicated that task completion time measures were robust.

Competing visual demands between eye fixation measures (IP glance duration and forward-viewing time) existed in both contexts. The results in eye fixation measures showed an inverse relationship between IP glance duration and forward-viewing time. This relationship was consistent with the findings of Kurokawa and Wierwille (1990). When the driving task became more difficult (higher speeds), shorter IP glance duration and longer forward-viewing time were observed.

Visual demands from the road scene were associated with the number of and radius of curves in the road, the field of view, and the sight distance (Hulse, Dingus, Fischer, and Wierwille, 1989; Green, Lin, and Bagian, 1993). Small radius curves and limited field of view resulted in a shorter sight distance in the simulator than on the road, because of geometric considerations. Since the visual attention demand was high in the simulator, drivers had to pay more attention to the forward scene, and could not afford to spend much time looking at the IP. Hence, the driving workload in the simulator could be assumed to be much greater than the driving workload on the road.

Driving performance showed poor absolute validity in the simulator. Lateral control measures (the standard deviation of steering angle and standardized lane position) showed that lane keeping was less precise in the simulator than on the road. These results were consistent with those obtained from previous studies (Blaauw, 1982; Reed and Green, 1995). The poor absolute validity in the simulator was due to three factors: the more complex (and unmatched) road geometry, the lack of risk in the simulator, and the absence of off-road cues. Since the degree of curvature of the simulated roads was larger than those on the road, keeping the vehicle centered in the lane was more difficult, and as a consequence, the standard deviation of the lateral control measures was much larger in the simulator than on the road. Since there was less risk involved in driving the simulator (a crash had minimal consequences), drivers were less strict in maintaining lane position. Moreover, the lack of off-road cues from the fixed-base driving simulator contributed to the large lateral excursions in the simulator. Drivers could go off the road without noticing such (particularly while attending to the secondary task). On the road, body and steering wheel shake, tire noise, and other cues would be immediately perceived. (Steering wheel shake and off-road sounds were added to the UMTRI simulator after this experiment was completed.) Since departures were less likely to be recognized immediately (and were more severe), recovery took longer in the simulator.

In addition to the lateral control measures, the longitudinal control measures (the standard deviation of throttle position and vehicle speed) showed poor absolute validity in the simulator. The simulated road was perfectly flat and there was no headwind variability. Consequently, drivers could maintain a constant speed by keeping their right foot on the accelerator in a fixed position. This was not the case outside the laboratory where all of the roads had some slight grade to them and the wind was never perfectly still. (These weaknesses were also overcome in later upgrades of the simulator.)

Driving performance showed good relative validity in the simulator. From the individual comparisons, results of the simulator showed a linear relationship between the three speed conditions and driving performance. This linear relationship indicated that as the driving speed increased in the simulator, the simulator became more difficult to control. Since driving speed was directly related to workload (high speeds were difficult to control and required more of the drivers' visual capacity), the simulator was sensitive to the workload of driving. Hence, varying levels of driving workload could be easily created in the simulator. This finding indicated that the simulator is an appropriate tool for evaluating secondary tasks, where workload is a critical issue.

Driving performance was expected to degrade when subjects drove and performed the secondary task simultaneously, since drivers needed to focus more on the secondary task instead of on driving. However, from the results, driving performance was worse on the road while driving without a secondary task than with secondary task. This result was inconsistent with that obtained by Reed and Green (1995). This unexpected result might be due to the road geometry selected for on-the-road driving. In the with-task condition on the road, only straight road sections were selected for safety reasons (it would be too dangerous for drivers to make a curve and perform the secondary task simultaneously with other vehicle around). However, in the without-task condition, driving performance data from both straight and curve sections were collected. Since keeping the vehicle on the center lane during a turn was harder to do than on a straight section, the driving performance in the without-task condition was worse, and lane keeping in the with-task condition was much easier since the road sections were straight.

The addition of a secondary task affected driving performance in both the simulator and on-the-road driving. From the results, the input measures (the standard deviation of steering angle and throttle position) were all greater in the without-task condition than in the with-task condition for both the simulator and on-the-road driving. This indicated that driving input was more constant while drivers performed secondary tasks. However, drivers' lane-keeping and speed-keeping performance might not necessarily be better in the with-task condition. The smaller deviation in input measures might be caused by the drivers' lack of visual attention on the driving task while they focused more on the secondary task. Since the drivers did not control the throttle as much while performing the secondary task, vehicle speed (a direct output of throttle manipulation) remained more constant in the with-task than the without-task condition. From the results, the standard deviation of steering angle for the simulator was smaller in the with-task condition, whereas the standard deviation of standardized lane position was greater in the same condition (lane position was a direct output of steering angle). This inverse relationship could be explained by the following driver behavior. Drivers tended not to steer as much while performing the secondary task. If the heading error was not very small, then not changing the steering would result in driving off the road, and hence the deviation in lane position. This problem was not apparent on the road, since drivers were more aware of the traffic (because of the risk of getting into crashes on the road), and their attention on steering was more acute. In addition, since data collected on the road for the with-task condition was all from straight sections, and that collected for the without-task condition contained both straight and curve sections, lane deviation was greater in the without-task than the with-task condition because staying at the center of the lane was more difficult.

The UMTRI driving simulator was concluded to be an appropriate context for evaluating secondary tasks. First, the simulator provided absolute safety that an experiment on the road could not achieve. In fact, no secondary task was given to drivers when driving on a curve on the road in this experiment due to safety reasons. Second, secondary task performance was robust as it was unaffected by the two different contexts. Hence, studies in secondary tasks could be run in either environment. In addition, the simulator was sensitive to driving performance changes related to workload. Since a linear relationship existed between driving performance and driving workload, different driving performance could be expected by changing the driving speed in the simulator. The 35 mi/hr condition in the simulator provided the closest match to the on-road conditions.

3. EXPERIMENT 2 - EFFECT OF PRACTICE ON DRIVER PERFORMANCE IN THE SIMULATOR

3.1 INTRODUCTION

Prior to conducting evaluations of secondary controls in a driving simulator, it was first necessary to gain a better understanding of the effects of practice. For a perceptual-motor task involving skill, the time to complete a task gradually decreases with practice. This, by definition, is learning. In dual-task contexts, one learns both how to perform the individual tasks and how to integrate them together (Wickens, 1992). However, it was uncertain how much secondary task and driving performance improved with practice in the driving simulator. Also uncertain was how the process of completing the secondary task changed with practice.

As is common knowledge, reinforced by the first experiment, differences in individual driving behavior and performance are quite large, and vary from situation to situation. For example, one person might drive a familiar vehicle every day for ten years, and remember how to operate every control (secondary task) without any additional visual references. In contrast, another person while driving an unfamiliar rental car in a strange city, may find turning on the radio baffling, even after numerous looks toward it. The difference between these two situations reflects both the differences between drivers and the effects of practice on task performance.

Completion of a secondary control task by unfamiliar drivers while driving can be thought of as a six step process.

1. Drivers locate the control grouping having the desired function. Visual landmarks need to be acquired.
2. Drivers find the control to operate. Included in this step may be determining the distance to move, planning the movement to the control, determining how to grasp it, and based on direction-of-motion stereotypes and control affordances, determining how to move the control.
3. Drivers move their hands to the control. Drivers usually move their hands toward a desired control without any visual references.
4. Drivers look at the control for final guidance of their hand to the control and grasping the control head. Practice might change this use of visual reference to tactile reference, where drivers only need to feel the control, rather than look at it to grasp it.
5. Drivers actuate the control.
6. Drivers monitor tactile, auditory, and sometimes visual feedback to determine when the operation is completed.

Drivers familiar with a vehicle perform secondary tasks quite differently from unfamiliar drivers. In step 1, location is retrieved from memory instead of being visually acquired. In step 2, active planning may be minimal, with the driver instead relying upon a stored motor program. While step 3 may be relatively unchanged, step 4, final guidance, tends to rely much more on tactile feedback in familiar drivers. Steps 5 and 6 are relatively unchanged, except in step 6, where the use of visual feedback is reduced. In brief, the task is highly automated in familiar drivers.

Since the safety and usability of secondary functions to both first time users and experienced users is of interest, and task execution could change in fundamental ways with practice, practice effects were examined in detail in the second experiment. This experiment was conducted entirely in the driving simulator.

In light of the previous discussion, the following issues were addressed:

1. How does driving task and secondary task performance improve with practice?
2. How does the structure of secondary task execution change with practice?
3. Do the effects of practice interact with the difficulty of the driving task?

This last issue has implications for the workload level that should be selected in driving simulator evaluations of secondary tasks.

3.2 METHODS

In this experiment, the effects of practice on an in-vehicle secondary task performed in the UMTRI driving simulator were examined. Subjects drove the simulator for 20 consecutive work days. Driver performance, both primary driving and secondary task measures for different speed and secondary task conditions was considered.

3.2.1 Subjects

For reasons of convenience, the first two authors of this report (a man and a woman) served as subjects. They were 26 and 22 years old respectively, and had little prior experience with the UMTRI driving simulator.

3.2.2 Driving Simulator and Test Route

The driving simulator and road driven were identical to that used for the simulator portion of experiment 1 (simulator validation). Only two speed conditions were examined (35 and 50 mi/hr). An example of a typical simulated road scene appears in Figure 2.2.3.

3.2.3 Experimental Design and Procedure

As shown in Table 3.2.1, there were three independent variables. Each subject participated for 20 consecutive days (except weekends and holidays), at approximately the same time of day, though the simulator schedule for other studies forced some exceptions. The three test conditions chosen represented an economical means to examine speed (workload) and secondary task effects. The fourth condition, 35 mi/hr without a secondary task, was eliminated to reduce the total running time for both subjects to fit within the constraints of simulator availability.

Table 3.2.1. Independent variables in learning experiment

Independent variable	Levels
Subject (S)	male and female
Day (D)	1 - 20 days
Condition (C)	(a) 35 mi/hr with secondary task
	(b) 50 mi/hr with secondary task
	(c) 50 mi/hr without secondary task

Each test block took approximately 5 minutes, with each subject taking about 20 minutes to complete each daily session. (See Figure 3.2.1.) The order of the three conditions was partially counterbalanced across subjects and days. Subjects were instructed to maintain speed and complete the secondary task as quickly as possible.

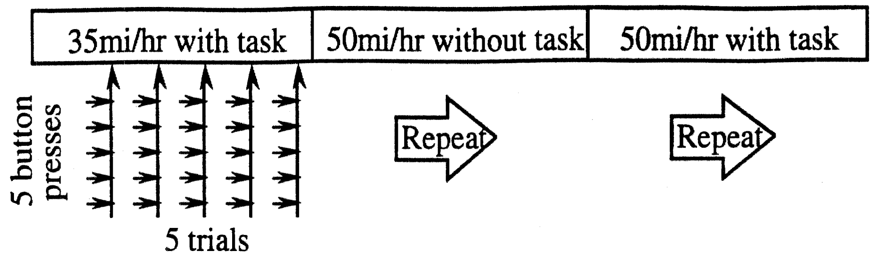


Figure 3.2.1. Procedure of the learning experiment

As can be seen in Table 3.2.2, all dependent measures in the driving task were identical to those in the simulator validation experiment except that eye fixation data was not analyzed since the emphasis was on performance data (and to save time).

Table 3.2.2. Dependent measures in learning experiment

Task category	Data source	Dependent measure
Secondary task	hand movements	interbutton response time (s)
		task completion time (s)
Driving	lateral control	s.d of steering angle (deg)
		s.d of standardized lane position (%)
	longitudinal control	s.d of throttle position (%)
		s.d of vehicle speed (mi/hr)

3.3 RESULTS

Prior to the main analysis, the histograms of the acquired data was compared with those of the log (1+x) transformed data to verify the normality of distributions prior to ANOVA tests. The distributions of both the secondary task and driving performance tended to be lognormal, though the departure from normality was not extreme. Hence, use of ANOVA was appropriate.

Accordingly, a fixed effects ANOVA was used to investigate the effects of the independent variables on secondary task and driving performance. As before, the criterion for statistical significance was $p<0.05$. The levels of the condition variable examined depended on the dependent measure. For in-vehicle task completion measures, only two conditions were examined (35 mi/hr with secondary task and 50 mi/hr with secondary task). For lateral and longitudinal control, all three conditions were considered.

3.3.1 Secondary Task Performance

From the ANOVA summary (Table 3.3.1), the results indicated significant effects of day ($p<0.01$) on interbutton RT (Figure 3.3.1) and task completion time (Figure 3.3.2). These results showed significant performance shifts across the 20 consecutive days. A Scheffe's test showed significant time decrement in hand movement measures across the first four to five days. However, the size of decrement decreased over time though decrements continued to occur throughout the experiment. Also note that other than individual differences, there were no interactions with day, suggesting that initial performance would be indicative of long-term (practiced) performance.

Figure 3.3.1. ANOVA summary of learning experiment

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Subject (S)	++	++	n/a	n/a	n/a	++	++		
Condition(C)			n/a	n/a	n/a	++	++		
Day (D)	++	++	n/a	n/a	n/a	++	++	++	
S x C			n/a	n/a	n/a	++	++		
S x D	++	++	n/a	n/a	n/a			+	+
C x D			n/a	n/a	n/a	+			

(++: $p < 0.01$; +: $p < 0.05$)

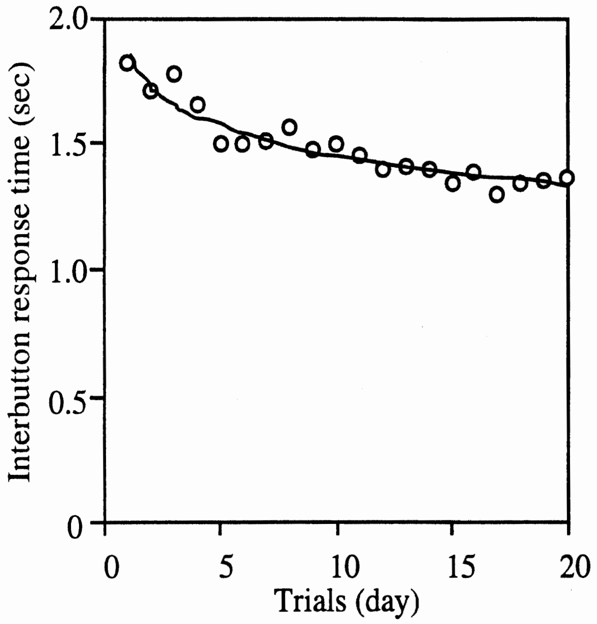


Figure 3.3.1. Effect of day on interbutton RT

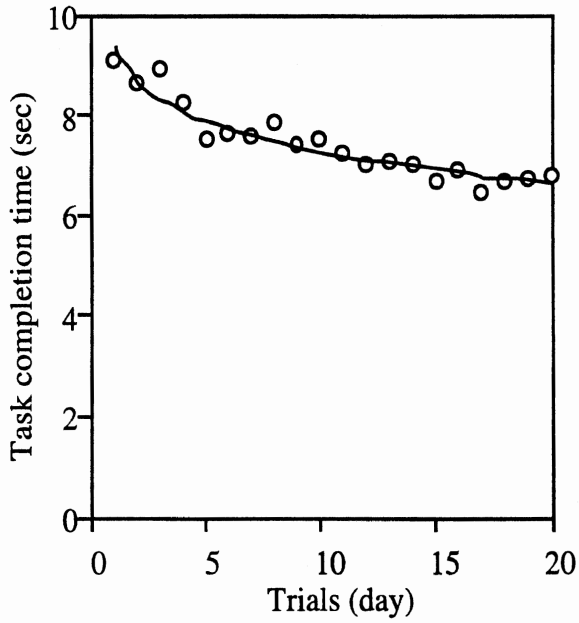


Figure 3.3.2. Effect of day on task completion time

A common finding in studies of perceptual-motor tasks is that task completion time decreases with practice, usually in an exponential manner. Buck and Cheng (1993) compared learning curves of power and exponential models for empirical data. The power model has been formalized as the Power Law of Practice and is shown below. Buck and Cheng report that power models provide greater simplicity, but sometimes exponential models provide better fits to human learning data. The Power Law postulates that the rate of time decrement is approximately proportional to a power of the amount of practice. The learning curve can be approximated by:

$$T_N = T_1 \; N^{-\alpha}$$

where

T_N = estimated performance time on the Nth block of trials
T_1 = estimated performance time on the first block of trials
N = trial block number, and
α = an empirically determined constant.

In this case, both models were fitted to the learning curves of interbutton RT and task completion time. The results from curve fitting (Table 3.3.2) showed much of the variance was accounted for

(r^2>0.8) for both task time measures. The values of the empirically determined constant were similar for interbutton RT and task completion time, suggesting the underlying processes were similar. Differences between models in the quality of fit were small.

Table 3.3.2. Learning curve fit (power and exponential)

Power	Learning curve equation	r^2
Interbutton RT	T_N=1.816 $N^{-0.10}$	0.895
Task completion time	T_N=9.078 $N^{-0.10}$	0.898
Exponential		
Interbutton RT	y = 1.733 exp (-0.015)t	0.846
Task completion time	y = 8.699 exp (-0.016)T	0.856

3.3.2 Driving Performance

From the ANOVA summary (Table 3.3.1), the results indicated significant effects of condition (p<0.01) and day (p<0.01) on the S.D. of steering angle (Figure 3.3.3) and the S.D. of standardized lane position (Figure 3.3.4). As a result of Scheffe's tests, significant differences in condition (p< 0.01) were found between 35 mi/hr and 50 mi/hr with-task, and between 35 mi/hr and 50 mi/hr without-task. However, a regression analysis showed small coefficients of determination (r-square: less than 0.3) within trials. These results showed that the lateral control measures were primarily consistent across test days. Therefore, the lateral control measures seemed to be primarily influenced by speed conditions and secondarily influenced by the secondary task and the effect of practice.

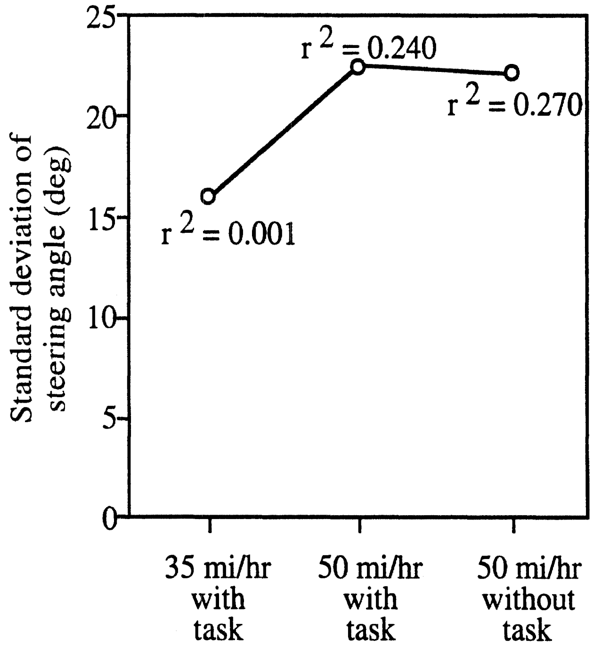


Figure 3.3.3. Effect of condition on S.D. of steering angle

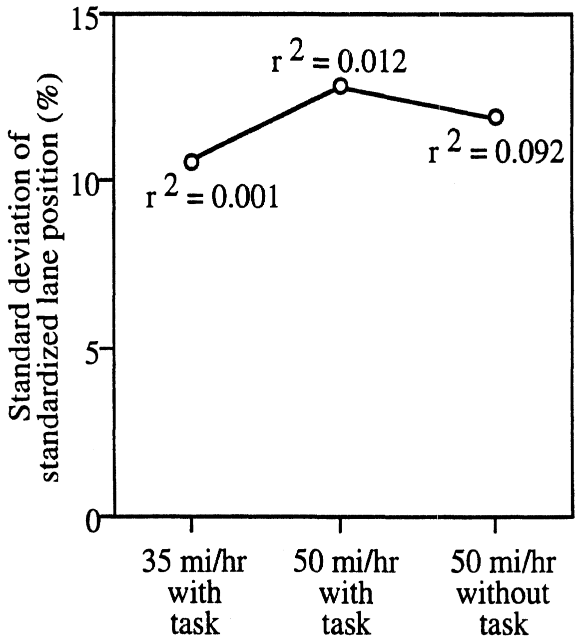


Figure 3.3.4. Effect of condition on S.D. of standardized lane position

The results indicated that the longitudinal control measures were inconsistent across days and no significant effects were found (Figure 3.3.5 and Figure 3.3.6).

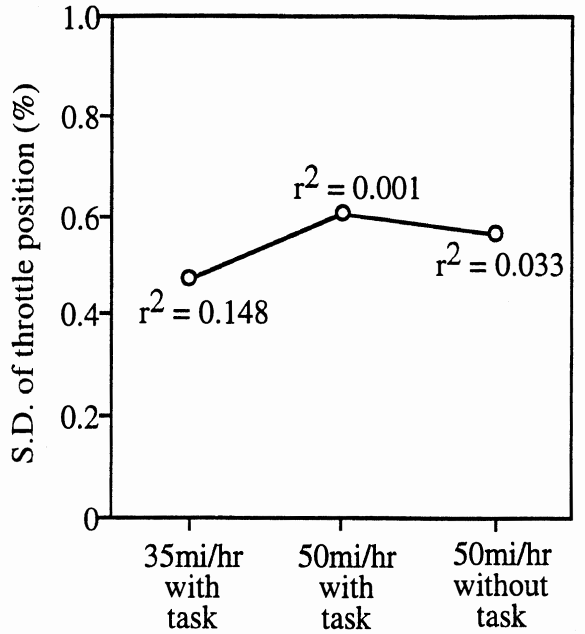


Figure 3.3.5. Effect of condition on S.D. of throttle position

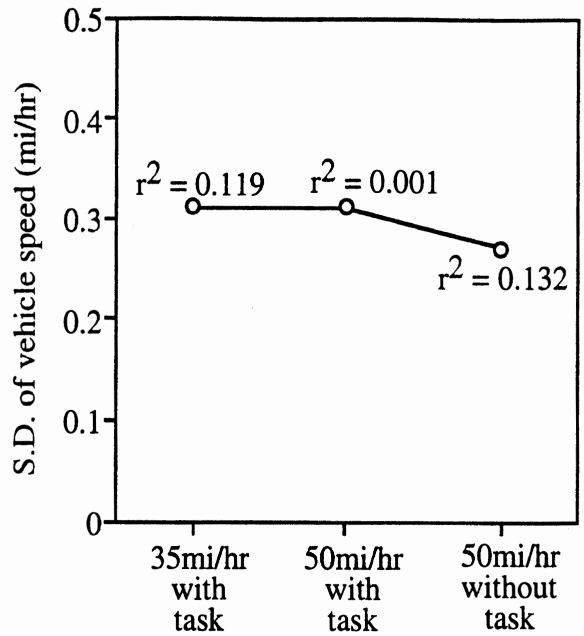


Figure 3.3.6. Effect of condition on S.D. of vehicle speed

3.4 DISCUSSION

The secondary task performance measures (interbutton RT and task completion time) showed significant performance shifts along the 20 days of the experiment. The results showed that the time to respond by pressing a button and to complete the secondary task decreased with practice. The performance time decrement was substantial in the first four to five days but continued to gradually decrease even on the last day of the experiment. The values of the exponent for the learning curves for interbutton RT and task completion time were the same.

The lateral control measures of driving performance (the S.D.'s of steering angle and standardized lane position) showed significant differences between the two speed conditions (35 and 50 mi/hr) and no significant difference between the presence and absence of secondary tasks. In addition, measures of longitudinal control of driving performance (the S.D.'s of throttle position and vehicle speed) showed no significant differences in speed conditions and in the presence and absence of secondary tasks. These results were so inconsistent across days that no significant effects of practice could be examined for driving performance. One way to explain the inconsistency in driving performance was that, since the simulator lacked any risk, some people drove in a more erratic manner. Hence, driving performance in the simulator seemed to be dominated by speed and less influenced by practice and the presence and absence of secondary tasks.

To conclude, practice affected secondary task performance, not driving performance. As mentioned earlier, if driving performance in the simulator improved with practice, then secondary task performance in the simulator might not be a true indicator of such performance in a real vehicle. However, if driving performance remained constant after practice, then the changes in secondary task performance in the simulator could reflect what would happen in a real vehicle, since driving performance in a real vehicle is constant.

4. EXPERIMENT 3 - USABILITY OF HIERARCHICAL MENU INTERFACES

4.1 INTRODUCTION

In terms of the intrusiveness of secondary tasks on driving, Kurokawa and Wierwille (1990) examined the effect of random crosswind on secondary task performance in a moving-base driving simulator. The results indicated that the strength of random crosswind influenced performance by increasing forward-viewing time while decreasing IP glance duration. Competing visual attentional demands existed between the IP and the forward scene. Zwahlen (1993) conducted a study to evaluate push button arrangements in automobiles. He concluded that the middle location on the IP produced slightly more accurate pushing performance than either the top or bottom locations. These results indicated the importance of relevant hand reach distance on completing secondary tasks.

More generally, minimizing the (1) visual and attentional demands and (2) the intrusiveness of secondary tasks on driving are essential criteria for driving safety. Given that the validity of the simulator had been established in the first experiment and that the second experiment established that minimal practice was required, this experiment examined the use of the driving simulator for the evaluation of menu-based secondary controls. This experiment focused on multipurpose controls and hierarchical menu structures. Driving and secondary task performance data were obtained as well as subjective ratings of drivers' preferences for controls and displays. In light of the previous research, the following issues were addressed in this experiment:

1. Which multipurpose control-hierarchical menu combination minimized visual-attentional demands of the secondary task?

2. Which multipurpose control-hierarchical menu combination minimized obstructiveness of secondary task on driving performance?

3. Which multipurpose control-hierarchical menu combination was the most preferred by subjects?

4.2 METHODS

In this experiment, all possible combinations of 4 controls (number keypad, rotary knob, remote trackball, and touchpad) and 3 hierarchical menu structures (2 x 6, 4 x 3, 8 x 2) were examined. Subjects drove the simulator while periodically performing secondary tasks using combinations of one specific control and one particular menu structure. Driver performance (driving and secondary task) and subjective ratings were obtained.

4.2.1 Subjects

A total of 16 licensed drivers (eight men and eight women) participated in this experiment. Each age group (young: age 20 to 30; old: 65 and over) consisted of 8 subjects (4 men, 4 women). All subjects were experienced drivers, but none of them were familiar with the UMTRI driving simulator.

4.2.2 Test Route

The same road selected for the simulator portion of the context comparison experiment was used in this experiment. Only two speed conditions were examined (35 and 50 mi/hr). An example of a typical road scene of the simulator can be found in the simulator validation section of this report (experiment 1).

4.2.3 Equipment and Materials

To simulate a typical in-vehicle secondary task, a number selection task was carried out in the UMTRI driving simulator. The task required the use of one of the four controls and a display. The equipment consisted of a Hitachi 5-inch LCD, four input devices (number keypad, rotary knob, remote trackball, and touchpad) that could appear in future vehicles, and a Macintosh Duo 270C portable computer which generated the secondary task program. The display was located at the top of the center console. The number keypad (Kensington Notebook) was located at the bottom of the center console. A custom-made rotary knob was located on the steering column, behind the right side of the steering wheel. The remote trackball (Microspeed MicroTRAC) was held by the subject when driving, so no permanent location was assigned. The touchpad (ALPS GlidePoint) was located at the button of the center console to the right of where the shift lever would be located (Figure 4.2.1). Because they were the most likely places for control and display elements in production vehicles, these locations were chosen as an attempt to maximize practical applications of the results.

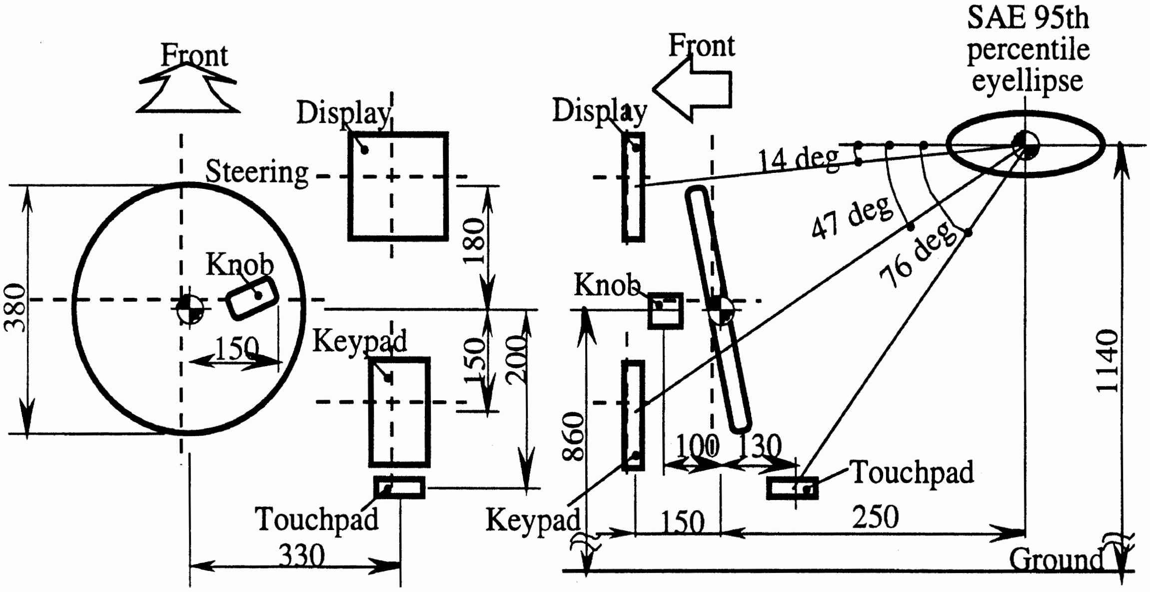
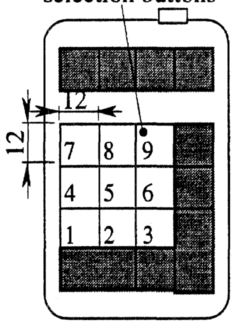
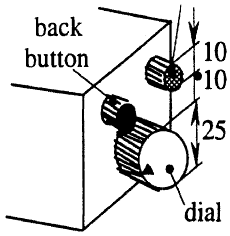
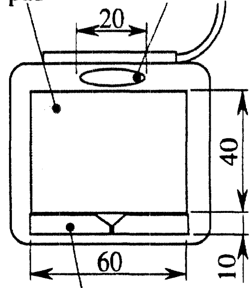


Figure 4.2.1. Location of the LCD and four multipurpose controls

The controls selected (number keypad, remote trackball, rotary knob and touchpad) differed in terms of how they were operated, how selection occurred, how visible the control was, and the ease of reaching the control (Table 4.2.1). Selection means was the method to select numbers on the display. The remote trackball and the touchpad were used to move a cursor to a location to select it. The cursor was free to move anywhere on the display. The square was chosen by pressing the select button on the devices. The rotary knob used a sequential selection method (12 positions (stops) per rotation). When the knob was turned one notch clockwise, the highlighted square on the display moved one square over in the same direction. Turning the knob counterclockwise had the opposite effect. The knob was to be rotated until the desired square was highlighted. Pressing the select button selected the square. The number keypad selected numbers directly. When a number was pressed on the keypad, the square on the display corresponded to the same location as the key on the keypad.

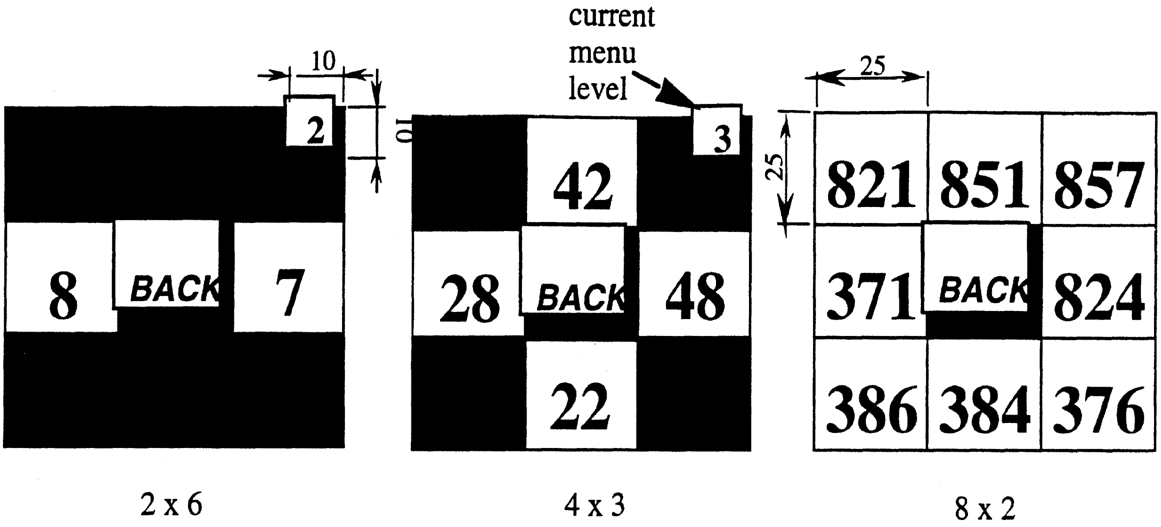
The selection procedure was determined by the number of steps to complete a selection task. The number keypad could complete the selection task with one step (press one button), whereas the remote trackball, touchpad, and rotary knob required two steps (move the cursor or highlight and press a button). The visibility and accessibility of the input devices were determined by their locations from the drivers and the obstructions in between.

Table 4.2.1. Controls used in the experiment

Control	Number keypad	Rotary knob	Remote trackball	Touchpad
Configu-ration	selection buttons  (Kensington)	selection button  (custom-made)		pad selection button  selection buttons (ALPS GlidePoint)
Operation	1. press the selection button corresponding to the location of the number displayed	1. rotate the dial until the target square was highlighted 2. press the selection button	1. hold control in the palm 2. put thumb on the trackball 3. roll the ball to move the cursor on the display 4. locate the cursor on a target square 5. press the selection button	1. put finger on the touchpad 2. drag the finger to move the cursor on the display 3. locate the cursor on a target square 4. press the selection button
Selection means	direct	sequential	infinite	infinite
Visibility	fair (ordinary IP location)	poor (behind the steering wheel)	good (possible to move anywhere)	poor (greater look-down angle)
Access-ibility	fair	good	good	good

4.2.4 Hierarchical Menu Structures

Three different menu structures (2 x 6, 4 x 3, 8 x 2) were examined in this experiment. They were classified by depth (the number of menu levels) and breadth (the number of choices per menu), shown in Figure 4.2.2. The design selected equalized the number of digits to be read.



(All dimensions are in mm.)

Figure 4.2.2. Display configurations for the three menu structures

A HyperCard program running on the portable computer generated numbers on the display. A series of 6 digit numbers were presented to subjects equalizing the number of digits to be read. In the 2 x 6 menu, 2 squares were displayed each time, with each square containing 1 digit. In the 4 x 3 menu, 4 squares were displayed each time, with each square containing 2 digits. In the 8 x 2 menu, 8 squares were displayed each time, with each square containing 3 digits. In addition, a BACK square in the middle of the display allowed subjects to go back one level in the menu structure if they made any incorrect selections. (See Figure 4.2.3 for an example entry sequence.) The back label did not appear at the top level, indicating the main menu.

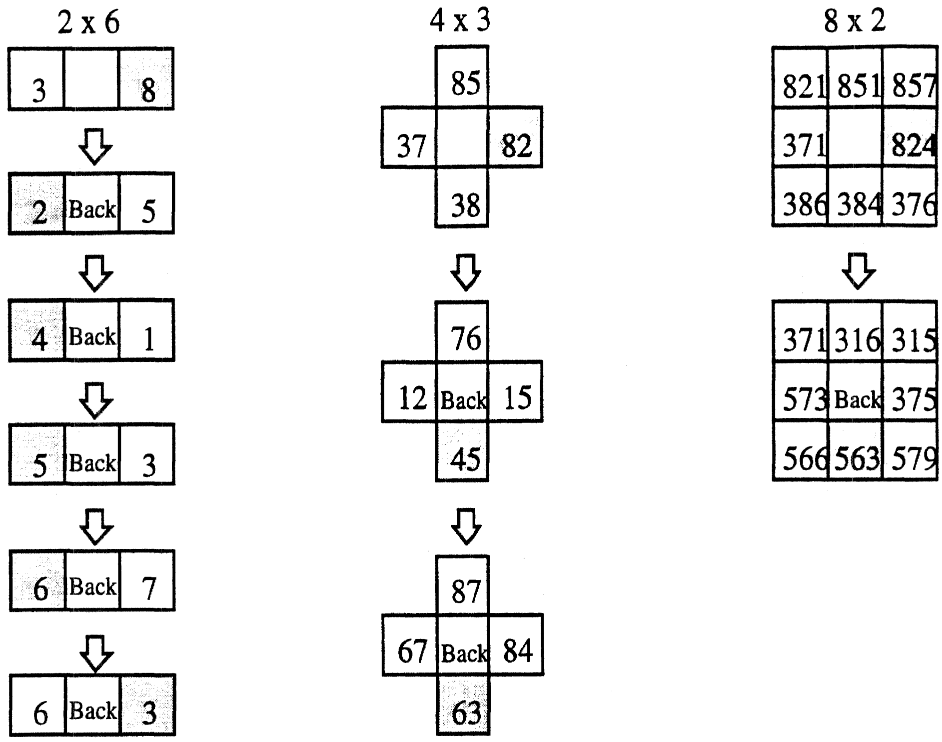


Figure 4.2.3. An example of selecting "824563" using different menu structures

4.2.4 Procedure

The experiment consisted of three parts: (1) a practice segment prior to data collection (learning to operate each of the four controls in static and dynamic driving simulator conditions), (2) a test segment at 35 mi/hr, and (3) a test segment at 50 mi/hr. Each part contained 4 blocks, one for each control. Each test block contained six trials, two trials for each menu structure. The orders of the speed conditions, controls and menu structures were counterbalanced across subjects (Figure 4.2.4).

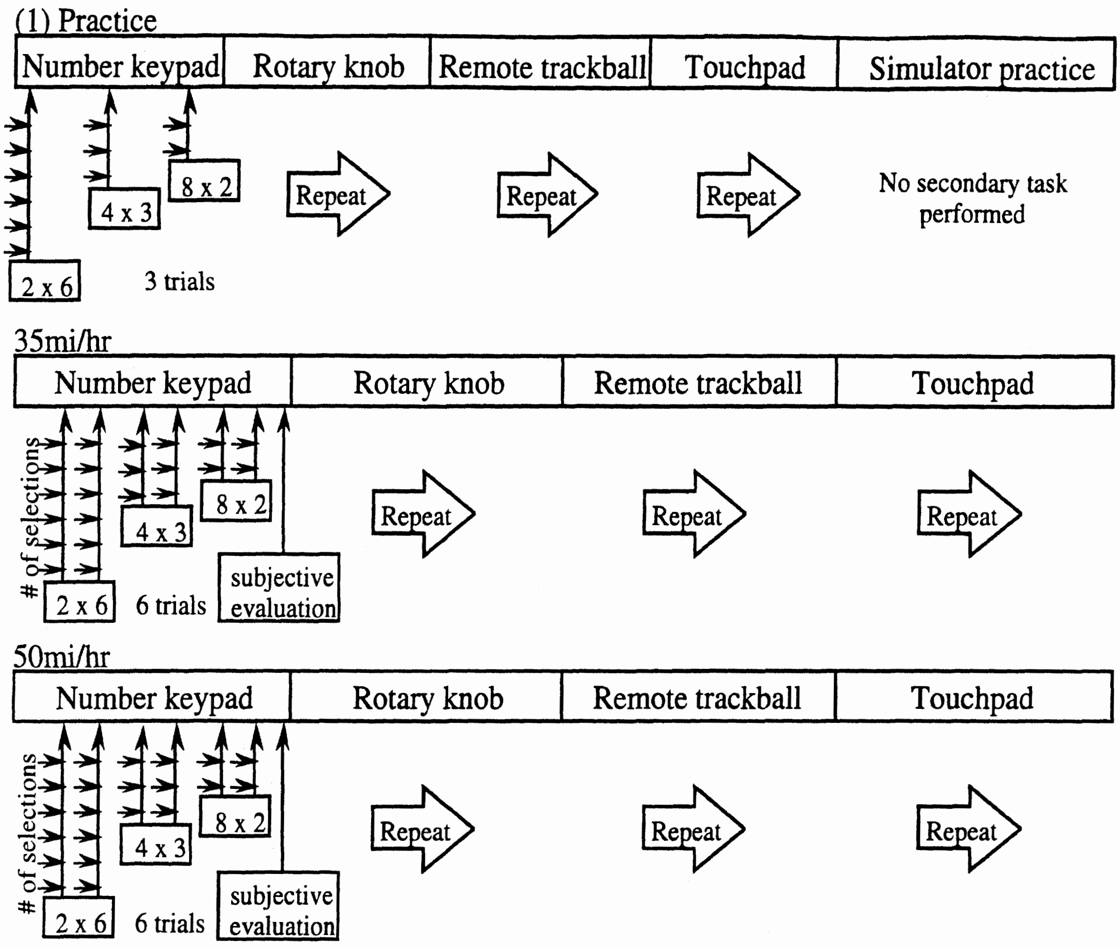


Figure 4.2.4. Procedure of the control comparison experiment

Depending upon the condition, subjects were instructed to maintain at constant speed (35 or 50 mi/hr) while driving the simulator. In each trial, subjects were verbally cued by a series of six numbers from the HyperCard program, and the identical numbers appeared at the upper left-hand corner of the screen. Once the auditory cue was finished, a specific menu structure appeared on the display, and subjects selected the numbers recited using a specific control while driving. An example using the 4 x 3 menu is shown in Figure 4.2.5.

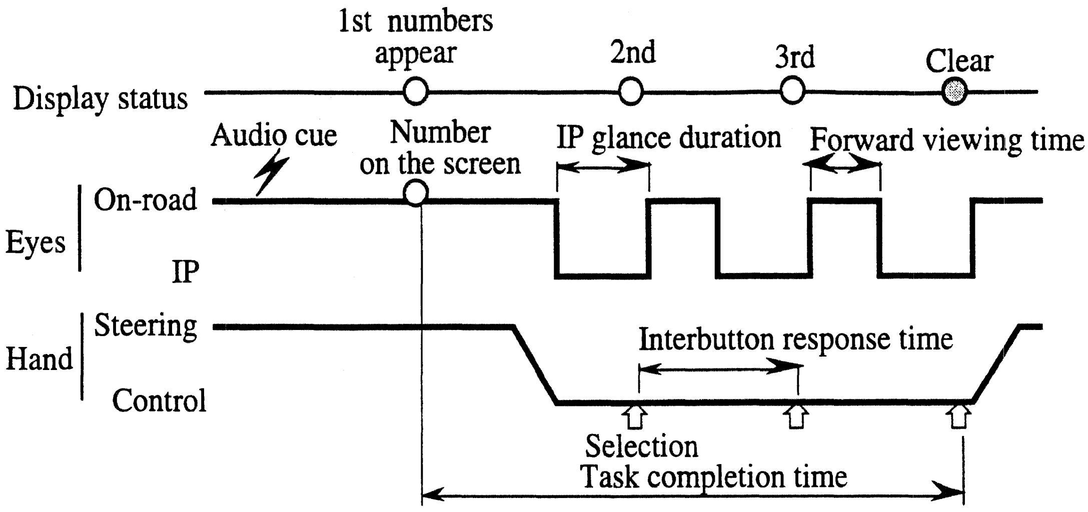


Figure 4.2.5. Mechanism of the secondary task and driver behavior (4 x 3 menu structure)

After each block, subjects completed a short evaluation to rate the different aspects of the menu structures and controls. After the last trial of the experiment, subjects completed a final evaluation. (See Appendix D.)

4.2.6 Experimental Design

Five independent variables were examined in the control interface experiment (Table 4.2.1). The within-subject variables were age and gender nested within subject. The between-subject variables were speed (2 levels), control (4 types), and menu (3 types) for a total of 384 responses/subject.

Table 4.2.1. Summary of independent variables

Independent variable	Levels
Age (A)	young, old
Gender (G)	men, women
Subject (S)	16 participants
Speed (Sp)	35, 50 mi/hr
Control (C)	number keypad, rotary knob, remote trackball, touchpad
Menu (M)	2 x 6, 4 x 3, 8 x 2

As can be seen in Table 4.2.2, all dependent measures in the driving task category were identical to those in the simulator validation experiment. Subjective ratings were examined to identify preferences for control type and menu structure.

Table 4.2.2. A summary of dependent measures

Task category	Data source	Dependent measure
Secondary task	hand movements	interbutton response time (sec)
		task completion time (sec)
	eye fixations	IP glance duration (sec)
		number of IP glances (time)
		forward-viewing time (sec)
	subjective evaluation	control type preference
		menu structure preference
		overall preference
Driving task	lateral controls	s.d. of steering angle (deg)
		s.d. of standardized lane position (%)
	longitudinal controls	s.d. of throttle position (%)
		s.d. of vehicle speed (mph)

4.3 RESULTS

An analysis of variance (ANOVA) test was used to investigate the effects of independent variables on secondary task and driving performance. The criterion set for statistical significance was $p<0.05$. The effect of menu on driving performance was not investigated, due to the constraints of experimental design (subjects operated all three menus by a control within a block). The histograms of the acquired data were compared to those of the log (1+x) transformed data to verify the normality of the distributions prior to the ANOVA tests. The distributions of secondary task (excluding number of IP glances) and driving performance tended to be lognormal and the distribution of the subjective ratings was normal (Figure 4.3.1). In contrast, the overall subjective ratings collected after the entire experiment were analyzed by non-parametric tests (Kruskal-Wallis and Wilcoxon rank-sum).

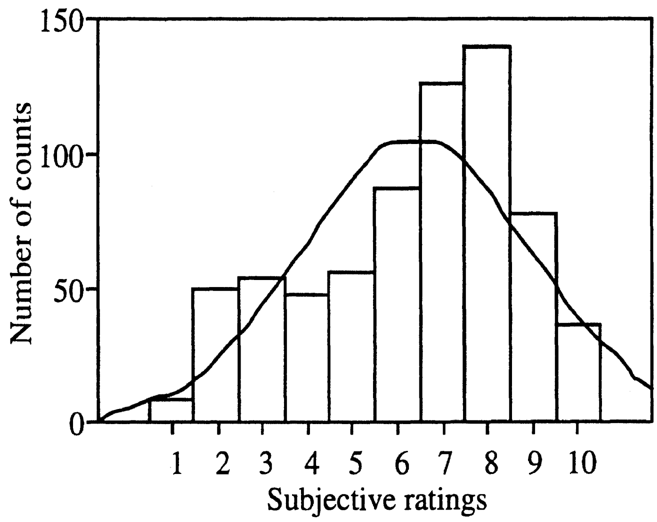


Figure 4.3.1. Distribution of subjective ratings

4.3.1 Secondary Task Performance

From the ANOVA summary (Table 4.3.1), the results indicated significant effects of control ($p<0.01$) on interbutton RT (Figure 4.3.2) and task completion time (Figure 4.3.3). A Scheffe's test on interbutton RT indicated significant differences in controls ($p < 0.01$) between the number keypad and other controls (rotary knob, remote trackball, touchpad), and between the rotary knob and two controls (remote trackball, touchpad). A Scheffe's test of task completion time indicated significant differences in controls ($p < 0.01$) between the number keypad and other controls (rotary knob, remote trackball, touchpad). These results suggested that the number keypad required less time to respond and to complete the same secondary task than other controls.

Table 4.3.1. ANOVA summary of evaluation experiment

	Hand movement		Eye fixation			Lateral		Longitudinal	
	inter RT	task comp	IP glance	for-ward	# of glances	steer-ing	lane pos	throttle angle	vehicle speed
Age (A)	++	++	++	++	++	++	++	++	++
Gender (G)		++	+	++					++
Subject (S)	++	++	++	++	++	++	++	++	++
Speed (Sp)	++	++	++	+		++		++	++
Control (C)	++	++	++		++	++			+
Menu (M)	++	++	++	++	++	n/a	n/a	n/a	n/a
A x G					+	++		++	
A x Sp			++						
A x C		+	++						
A x M						n/a	n/a	n/a	n/a
G x Sp			++					++	++
G x C	++			++					
G x M						n/a	n/a	n/a	n/a
S x Sp		++	++	+		+			
S x C		++	++	++	++	+			
S x M		+				n/a	n/a	n/a	n/a
Sp x C			++						
Sp x M						n/a	n/a	n/a	n/a
C x M	++	++	+	+		n/a	n/a	n/a	n/a

(++: $p < 0.01$; +: $p < 0.05$)

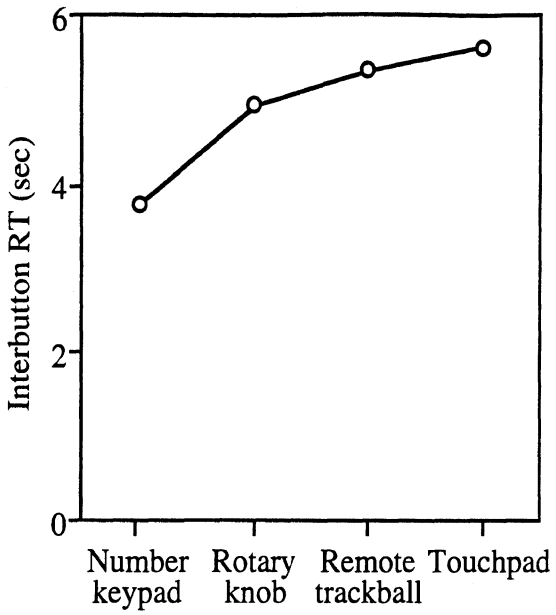


Figure 4.3.2. Effect of control on interbutton RT

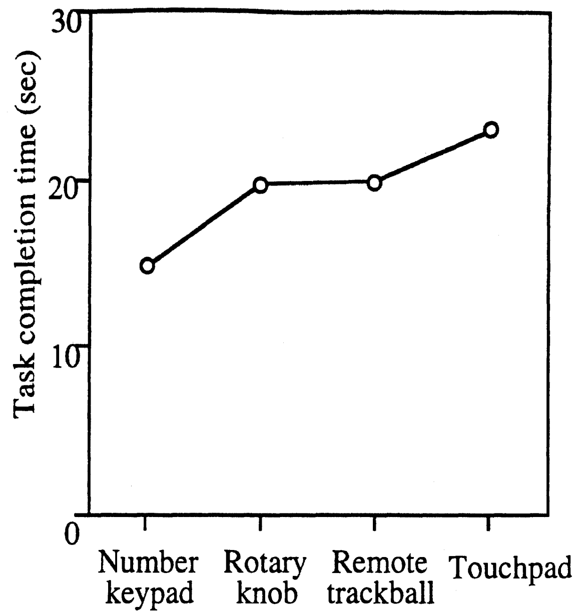


Figure 4.3.3. Effect of control on task completion time

The results indicated significant effects of menu ($p < 0.01$) on interbutton RT (Figure 4.3.4) and task completion time (Figure 4.3.5). A Scheffe's test of interbutton RT indicated significant differences in menus ($p < 0.01$) among all menu combinations. A Scheffe's test on task completion time indicated significant differences in menus ($p < 0.01$) between the 2 x 6 menu and other menus (4 x 3, 8 x 2), and a marginal difference ($p < 0.05$) between the 4 x 3 and 8 x 2 menus. These results suggested that interbutton RT increased with the number of selections (menu breadth), and task completion time decreased with the number of levels (menu depth).

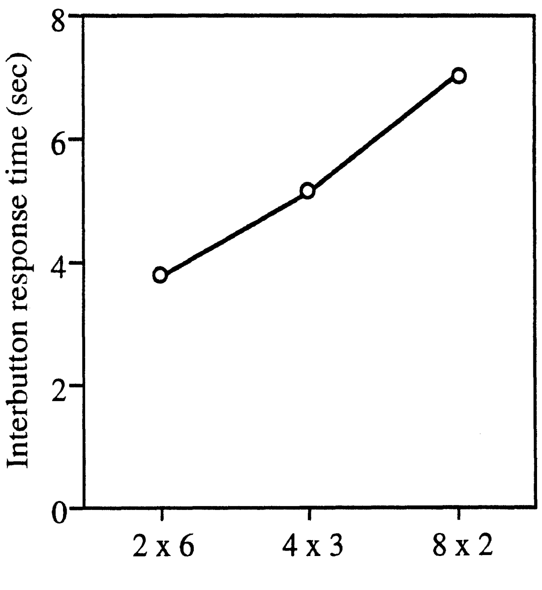


Figure 4.3.4. Effect of menu on interbutton RT

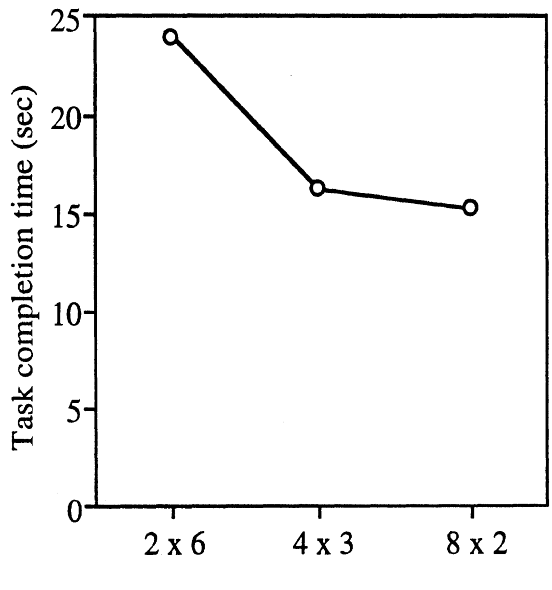


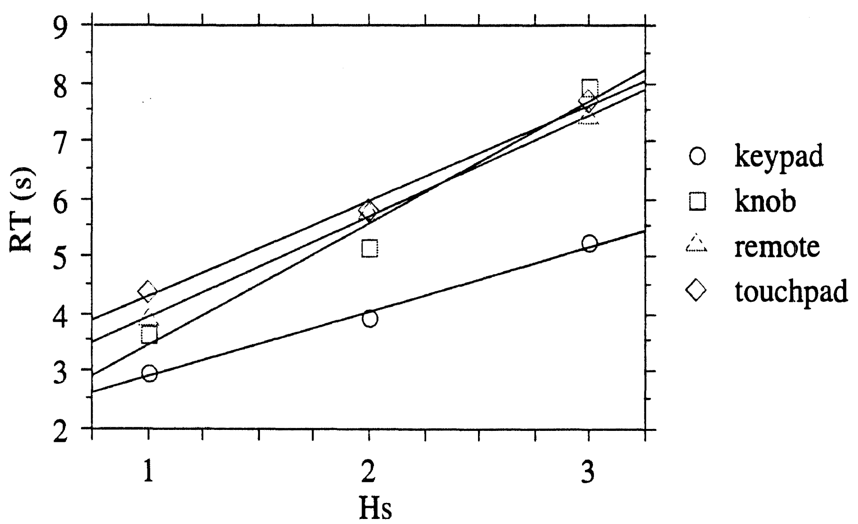
Figure 4.3.5. Effect of menu on task completion time

Hick and Hyman used information theory to quantify the uncertainty of stimulus events (Wickens, 1992). They found that choice RTs increased linearly with stimulus information as shown in the equation that follows. Here the Hick-Hyman Law (sometimes referred to as Hick's Law) was used to relate interbutton RT to the number of menu choices (menu breadth) at each level.

$$RT = a + b\,Hs$$

where RT = choice response time (sec)
a = constant (sum of processing latency unrelated to the reduction of uncertainty)
b = constant (added processing time/bit of stimulus information to be processed)
H = $\log_2 N$ (2 x 6 = 1, 4 x 3 = 2, 8 x 2 = 3)
N = number of alternatives

Given the interaction between control type and menu, as described below, separate Hick's Law equations were developed for each control type using the means for each control-menu combination. Figure 4.3.6 shows the relationships for all four controls along the r^2 values. Notice that the fits are highly linear. The percentage of the variance accounted for is at least 97 percent in all cases, 99 percent in all but one. Interestingly, the points for the three continuous controls (knob, remote trackball, and touchpad) were all grouped together and differences between them were small. Further, their times are consistently greater than the keypad, the only discrete control device. This hints at a general advantage of discrete selection devices for menus of this type.



keypad RT = 1.798 + 1.131 * Hs; R^2 = .993
knob RT = 1.322 + 2.127 * Hs; R^2 = .972
remote RT = 2.201 + 1.746 * Hs; R^2 = .999
touchpad RT = 2.667 + 1.642 * Hs; R^2 = .993

Figure 4.3.6. Hick-Hyman Law predictions for each control

When the predictions are made across control type, the r^2 value drops to 0.736, with the resulting equation (useful for making rough estimates) being RT = 1.997 + 1.622 * Hs. Thus, each additional bit of information adds 1.6 seconds with the minimum response time being 2 seconds for this task.

Further information on the relationship between controls and menus, a significant interaction ($p < 0.01$) for both interbutton RT (Figure 4.3.7) and task completion time (Figure 4.3.8), follows. Notice the consistent ordering of the controls for all menu structures except for the knob for the 8 x 2 menu for interbutton RTs, and that same menu when task completion time was considered.

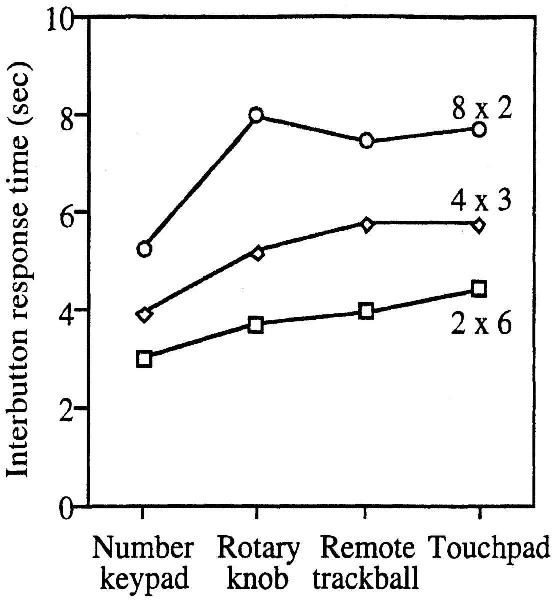


Figure 4.3.7. Control by menu interaction for interbutton RT

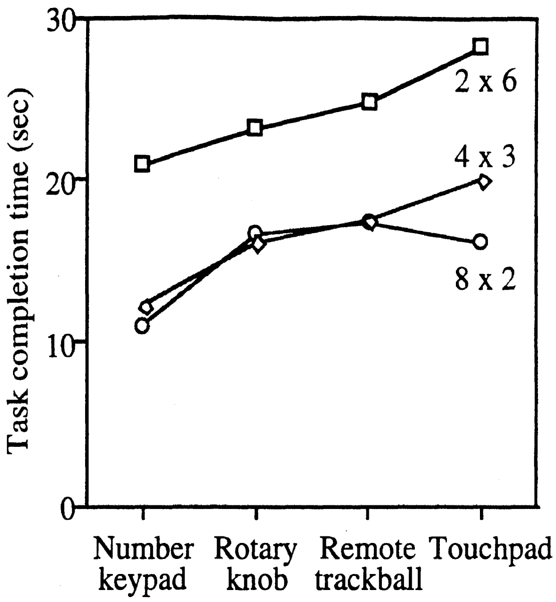


Figure 4.3.8. Control by menu interaction for task completion time

With regard to the glance data, the results were similar to the task time data. There was a significant effect of control ($p < 0.01$) on IP glance duration (Figure 4.3.9) and the number of IP glances (Figure 4.3.9). A Scheffe's test on IP glance duration showed significant differences ($p < 0.01$) in controls between the number keypad and two controls (remote trackball, touchpad), and between the rotary knob and the same two controls. In addition, a marginally significant difference ($p < 0.05$) was found between the remote trackball and the touchpad. A Scheffe's test of the number of IP glances showed a significant difference in controls ($p<0.01$) between the number keypad and the remote trackball, and a marginally significant difference ($p<0.05$) between the number keypad and the touchpad. These results suggested that the number keypad and rotary knob required a shorter IP glance duration and the number keypad showed fewer number of IP glances among all controls.

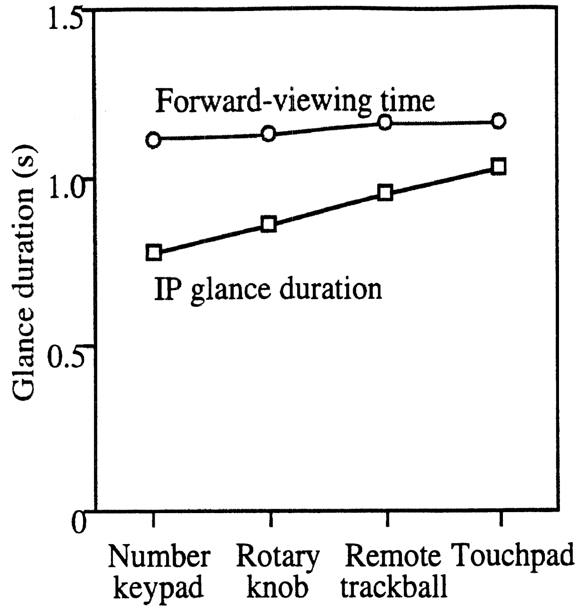


Figure 4.3.9. Effect of control on IP glance duration and forward-viewing time

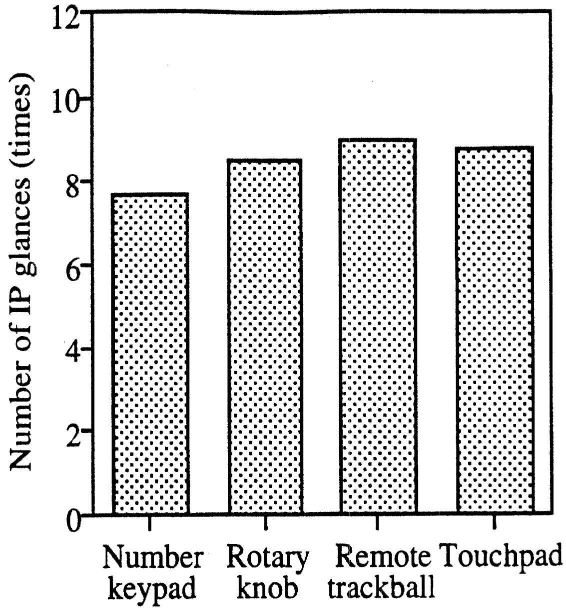


Figure 4.3.10. Effect of control on number of IP glances

The results indicated significant effects of menu ($p < 0.01$) on IP glance duration, forward-viewing time (Figure 4.3.11) and the number of IP glances (Figure 4.3.12). A Scheffe's test of IP glance duration and the number of IP glances showed significant differences in menus ($p < 0.01$) among all menu combinations. A Scheffe's test on forward-viewing time showed a significant difference ($p < 0.01$) between the 2 x 6 and 8 x 2 menus. These results showed an inverse relationship between IP glance duration and forward-viewing time.

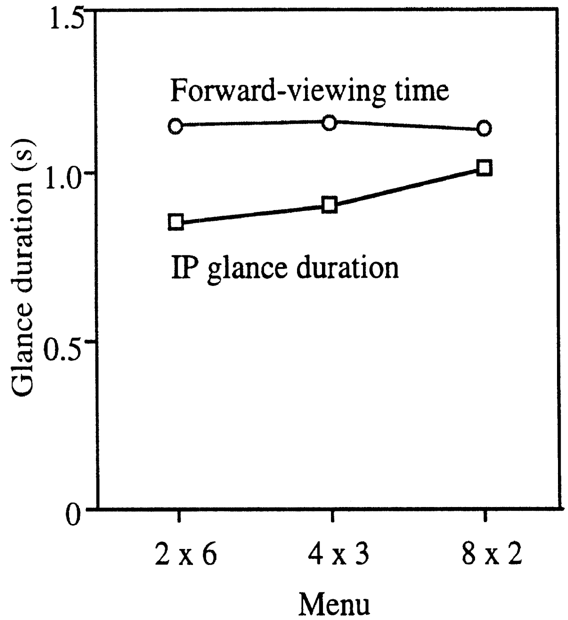


Figure 4.3.11. Effect of menu on IP glance duration and forward-viewing time

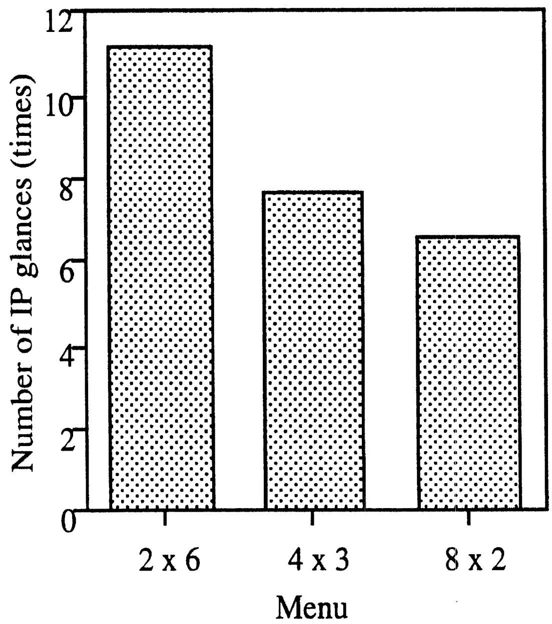


Figure 4.3.12. Effect of menu on number of IP glances

The number of IP glances decreased with the number of levels in the whole menu hierarchy (depth). In theory, the minimum number of IP glances should be identical to the number of menu depth hierarchy. For example, for the 2 x 6 menu, the drivers had to glance at the display at least six times to complete the secondary task. In this experiment, the ratios between the ideal and actual number of IP glances were calculated (Table 4.3.2). The table showed that the increased amount of displayed information (per menu) required more IP glances from drivers as expected. These results suggested that an inverse relationship between menu depth and the number of IP glances existed.

Table 4.3.2. Number of IP glances on each menu structure (per screen)

Menu	Mean number of IP glances (per screen)
2 x 6	1.87
4 x 3	2.55
8 x 2	3.27

In addition, the results indicated marginally significant control by menu interactions ($p < 0.05$) on IP glance duration (Figure 4.3.13) and forward-viewing time (Figure 4.3.14). A combination of the number keypad and the 2 x 6 menu required the shortest IP glance duration and the longest forward-viewing time.

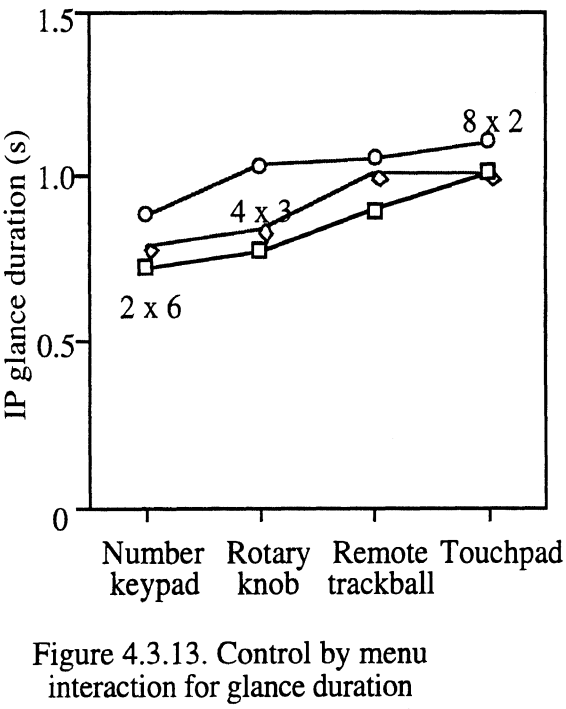


Figure 4.3.13. Control by menu interaction for glance duration

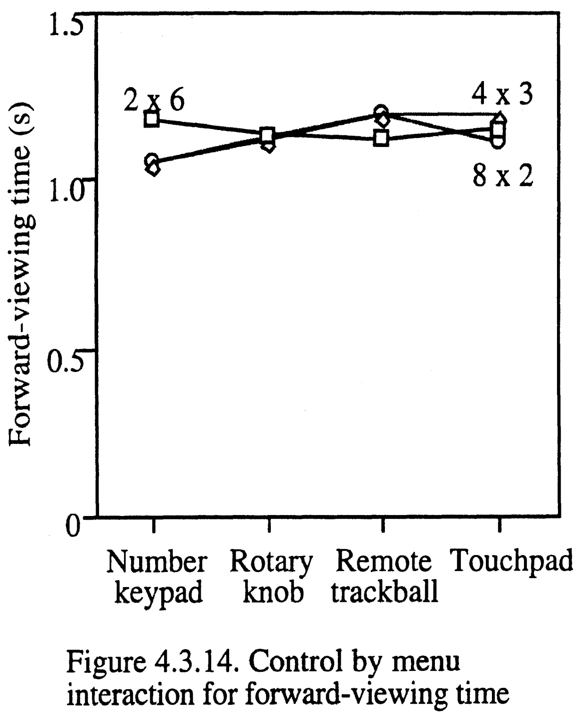


Figure 4.3.14. Control by menu interaction for forward-viewing time

The results indicated a significant age by speed interaction ($p < 0.01$) on IP glance duration (Figure 4.3.15) and a significant age by control interaction ($p < 0.01$) on IP glance duration (Figure 4.3.16). The young age group performed with a shorter IP glance duration than the old age group. Both young and old age groups performed with a shorter IP glance duration at the higher speed condition (50 mi/hr) than at the lower speed condition (35 mi/hr). These results suggested that speed conditions affected IP glance duration. The age by control interaction on IP glance duration showed that the difference between age groups was small when the number keypad was used. These results showed that the number keypad was effective for both young and old age groups to reduce IP glance duration.

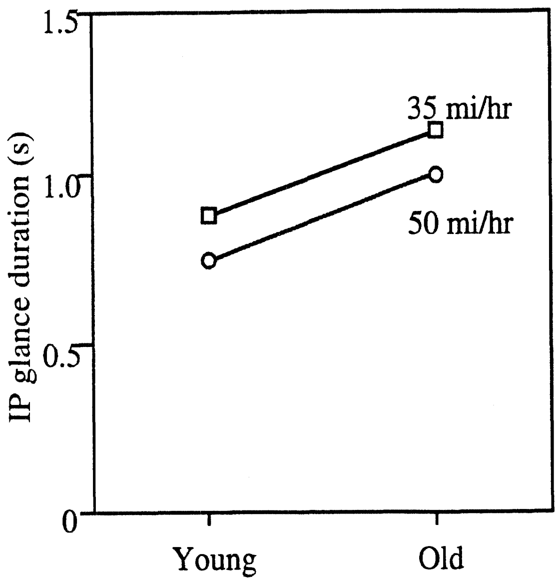


Figure 4.3.15. Age by speed interaction for IP glance duration

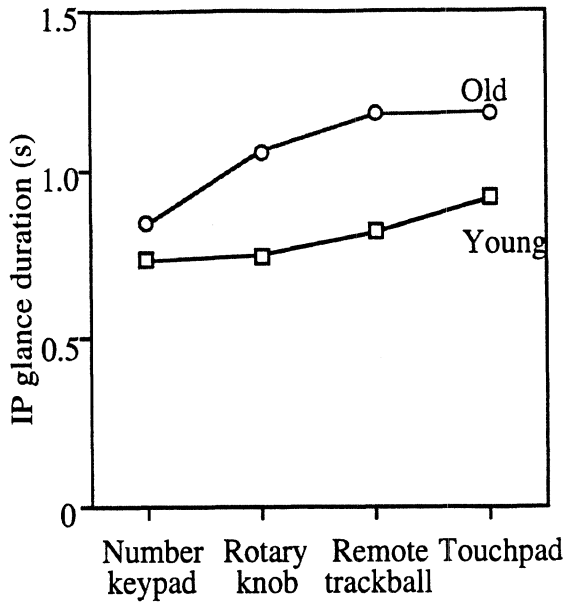


Figure 4.3.16. Age by control interaction for number of IP glances

4.3.2 Driving Performance

From ANOVA summary (Table 4.3.1), the results indicated a significant effect of control ($p<0.01$) on the S.D. of steering angle (Figure 4.3.17) and a marginally significant effect of control ($p<0.05$) on the S.D. of vehicle speed (Figure 4.3.18). A Scheffe's test of the S.D. of steering angle showed a marginally significant difference ($p< 0.05$) between the remote trackball and the touchpad. A Scheffe's test of the S.D. of vehicle speed showed no significant difference at this level. The remote trackball showed that the S.D. of steering angle was smaller than the touchpad. However, these results showed that the obstructiveness of secondary task on driving performance was small. Drivers tended to maintain the driving task and allocate limited resources to secondary task.

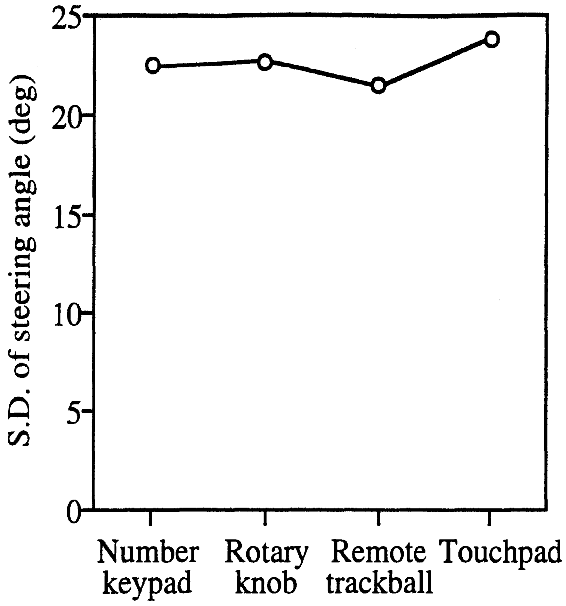


Figure 4.3.17. Effect of control on S.D. of steering angle

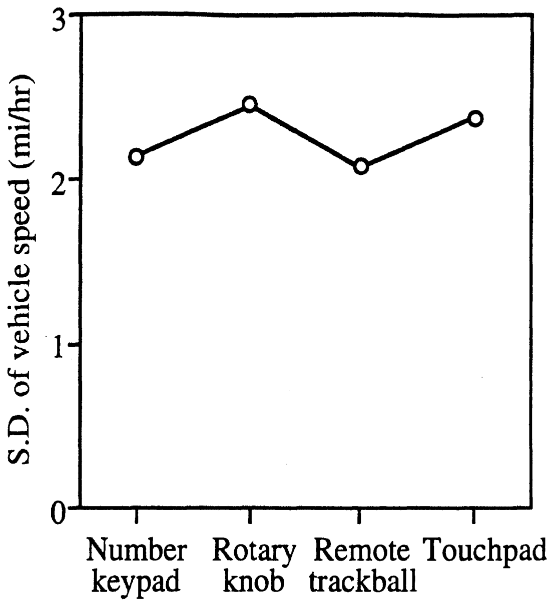


Figure 4.3.18. Effect of control on S.D. of vehicle speed

Differences in menu structure were not examined for the driving performance measures as the three menu structures were presented in a fixed order within blocks, and there was no clear demarcation in the driving data file as to when use of one menu structured ended and another began.

4.3.3 Post-Trial Subjective Ratings of Menu Usability

Two groups of ratings were obtained, ease of use (Appendix C, part 1) and driver preferences (Appendix C, part 2). The ease of use ratings--selection difficulty, accessibility of each control, difficulty of operating each control, and difficulty of driving the simulator--were collected after each menu-control combination test block. Each of these four dependent measures was analyzed separately. Data were utilized from all subjects who participated including one subject who did not complete all parts of the experiment. In several cases, the addition of partial data from subjects had a minor impact on the ratings obtained, not reflected in the statistical significance of outcomes.

Figure 4.3.19 shows the distribution of the selection difficulty ratings. In general, the ratings tend to reflect lower levels of difficulty (10=not difficult at all, 1=extremely difficult, mean=6.3).

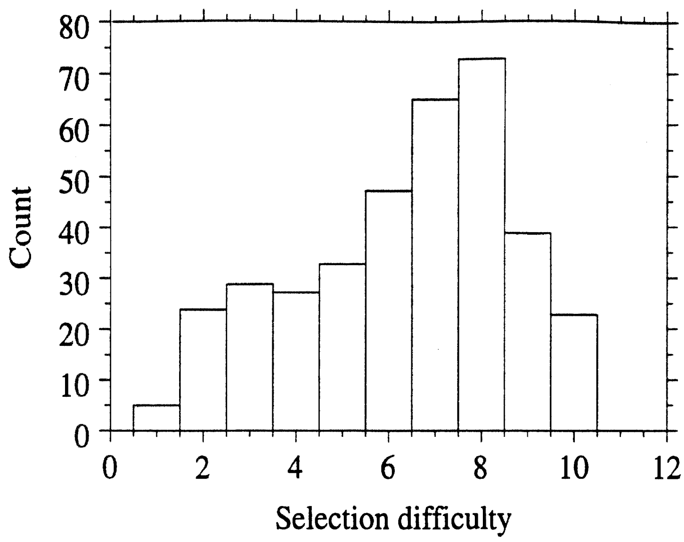


Figure 4.3.19. Selection difficulty ratings

Ratings of selection difficulty were examine in a partial factorial model involving subject factors (age, sex, the subject by sex interaction, and subjects main effects nested within age and sex), the device and condition factors (control, menu, and speed, as well as all combinations of them with each other), and interactions of age and sex with the three device and condition factors. Other factors were not included in the model as they were unlikely to be significant, and even if they were, would be difficult to interpret. Table 4.3.3 shows the resulting ANOVA summary.

Table 4.3.3. ANOVA summary of ratings of selection difficulty

Source	p	Source	p
Age (A)	+++	A x G	+++
Gender (G)	+++	A x Sp	
Subject (S) -nested (AxG)	+++	A x C	++
Speed (Sp)	+++	A x M	
Control (C)	+++	G x Sp	++
Menu (M)	+	G x C	
		G x M	
		C x M	
		Sp x M	
		Sp x C	
		Sp x M x C	

+++ = p<.001, ++ = p<.01

(Probability values of less than 0.05 are shown)

All of the main effects were significant. In general, young subjects offered higher ratings than older subjects (7.0 versus 5.5), and the ratings from men were slightly greater than those from women (6.4 versus 6.2). Figure 4.3.20 shows the age by sex interaction, one of the three significant interactions.

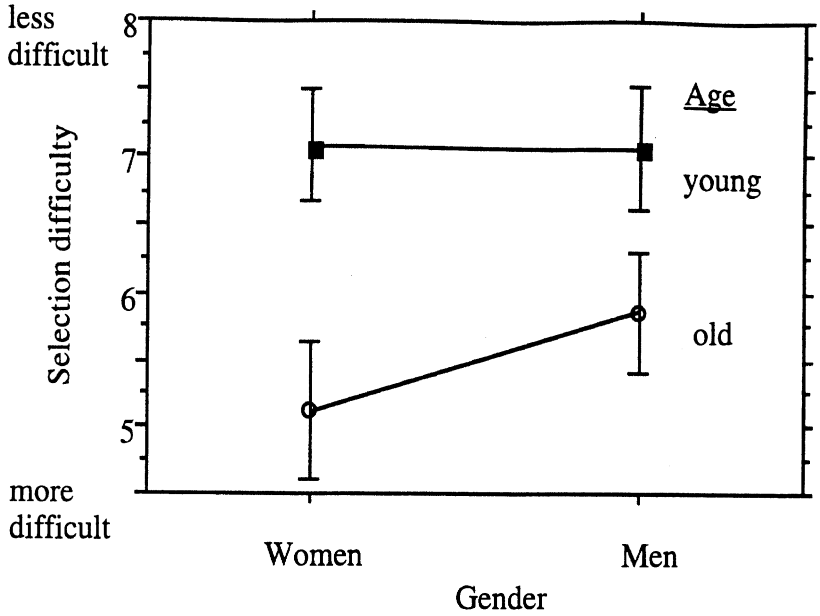


Figure 4.3.20. Effects of age and gender on ratings of selection difficulty

With regard to design factors, knobs were rated as less difficult to operate than keypads, trackballs, and keypads in that order (means of 6.9, 6.7, 6.1, and 5.6, respectively). (See Figure 4.3.21.) This pattern was consistent across all other factors except age (as indicated by the significant age by control interaction, Figure 4.3.22) with younger subjects favoring the knob and older subjects favoring the keypad. There is no evidence as to why this occurred. Of the menu structures, 2 x 6 was rated as least difficult (6.7) followed by 4 x 3 (6.4) and 8 x 2 (6.0). See Figure 4.3.21.

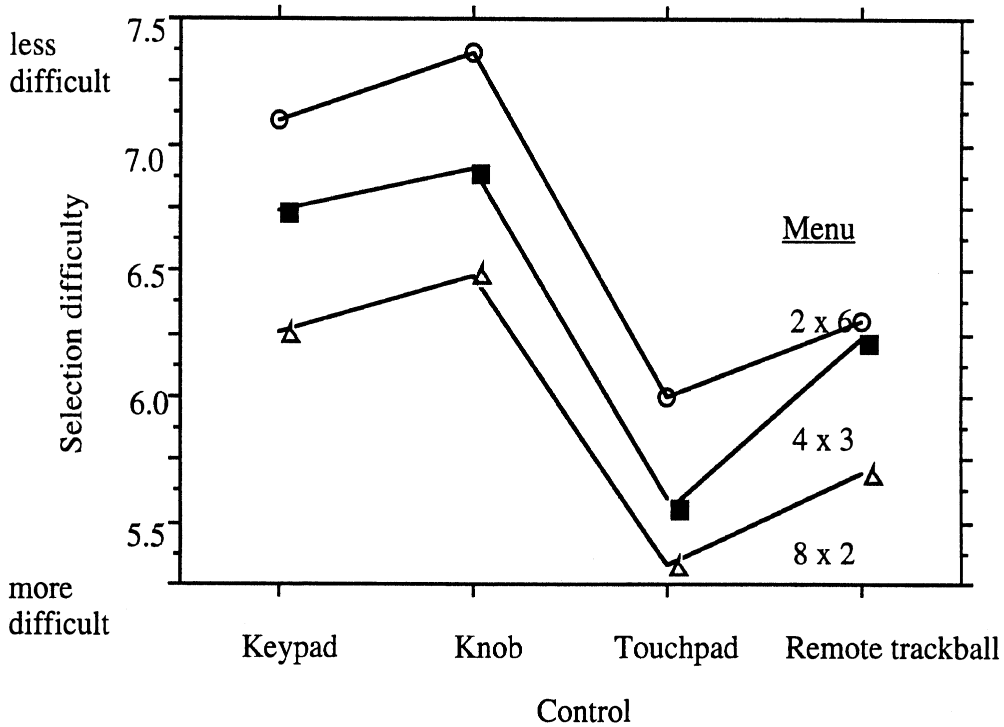


Figure 4.3.21. Effect of control type and menu structure on selection difficulty

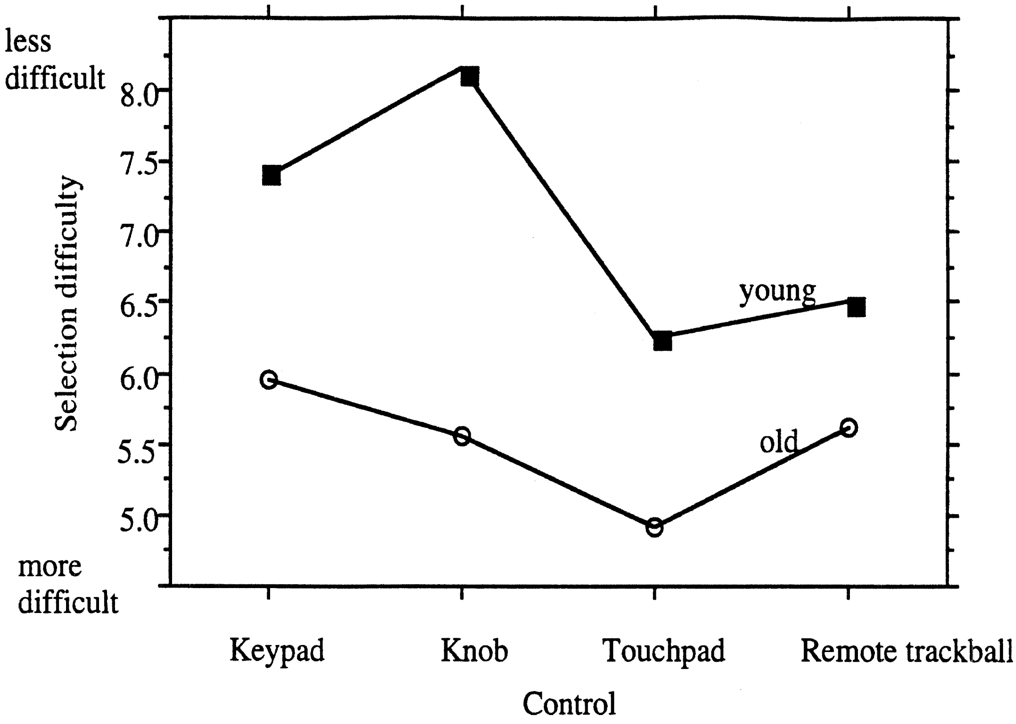


Figure 4.3.22. Age by control interaction for selection difficulty

Interestingly, as speed increased from 35 to 55 mi/hr, ratings of selection became more difficult (from 6.5 at 35 to 6.2 at 55). However, there were no interactions of interface design factors (control or menu) with speed, suggesting that in future studies testing at one speed should be adequate. (There was, as shown in Figure 4.3.23, an interaction with gender.)

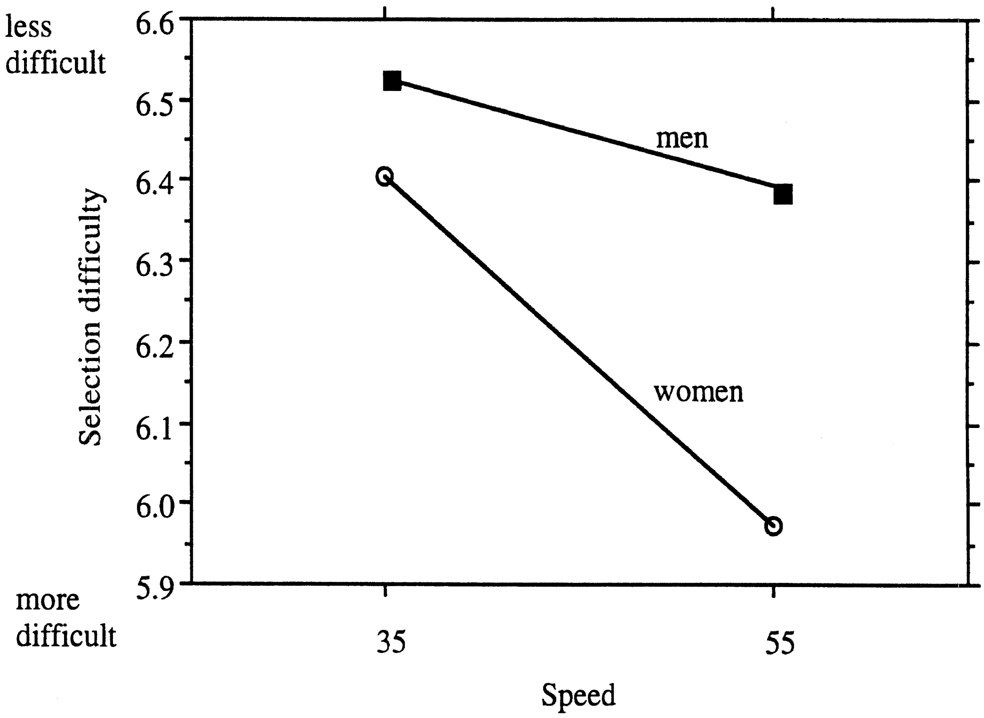


Figure 4.3.23. Speed by sex interaction for selection difficulty

4.3.4 Post-Trial Subjective Ratings of Control Usability

Questions 2 through 4 (accessibility, operation difficulty, and driving difficulty) of part 1 of the subjective evaluation were analyzed similarly, with age, sex, speed driven (35 or 50 mi/hr), and control (touchpad, keypad, trackball, and knob) being included in a full factorial model, and subjects used for replications. All factors were fixed effects. All interactions with more than two terms were pooled into the error variance as preliminary analysis showed they were not significant.

Figure 4.3.24 shows how accessible drivers reported the controls were (question 2). In that ANOVA there were significant differences due to age (p<.0002) and control (p<.0095) as well as age by gender interaction (p<.0345). No other effects were significant. On a 1 to 10 scale (1=hardly accessible, 10=easily accessed), younger drivers rated the controls much more accessible (mean=7.9) than older drivers (6.5). The knob was rated as slightly more accessible than the trackball/remote (7.9 versus 7.7), which was somewhat greater than the ratings for the touchpad and keypad (both 6.6). The age by sex interaction is due to the younger men and older women offering the lowest and highest ratings respectively, potentially reflecting their capabilities. (At younger ages, men generally perform better. In the older groups, the women usually perform better. Figure 4.3.25 and 4.3.26 show these significant effects.

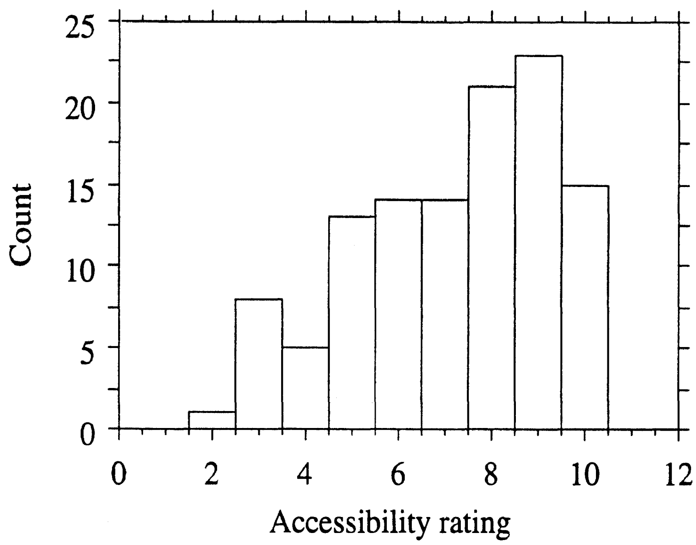


Figure 4.3.24. Control accessibility ratings

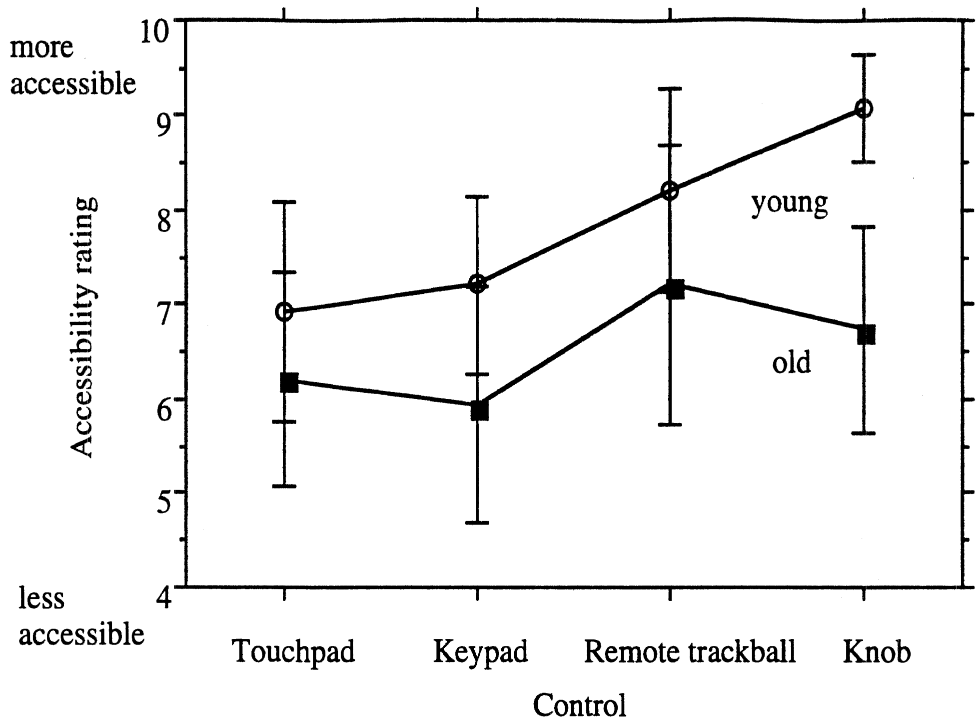


Figure 4.3.25. Effect of Control Type and Driver Age on Accessibility Ratings

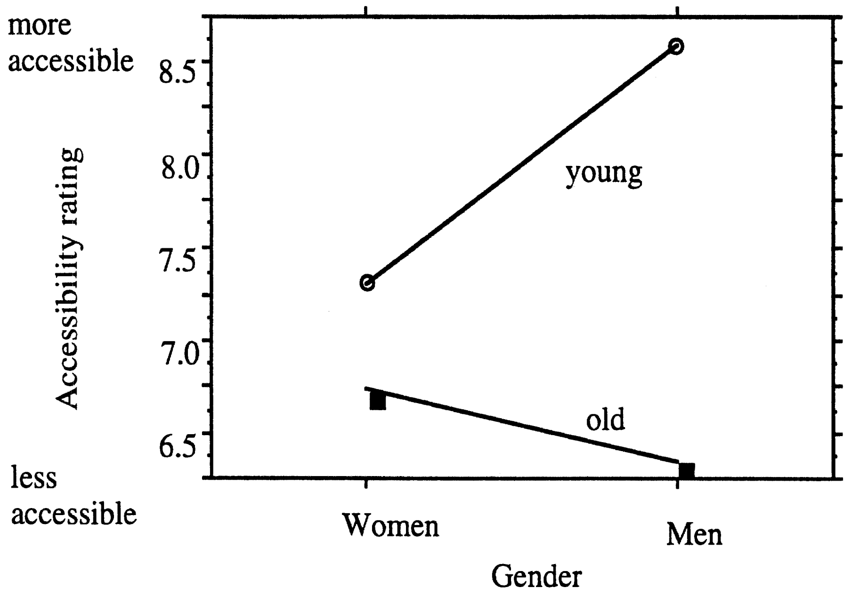


Figure 4.3.26. Effect of age and gender on accessibility ratings

Also examined were ratings of the difficulty of operating controls (question 3, part 1). Only the effects of age (p=.0088) and control (p=.0002) were significant. Ratings for young drivers (mean=7.5) were much greater than those for older drivers (6.5), where 1=extremely difficult to operate and 10=not difficult at all to operate. The two best rated controls were the keypad (mean=7.7) and the knob (7.6). Ratings for the trackball/remote (6.9) and touchpad (5.5) were much lower. Figure 4.3.27 shows the results.

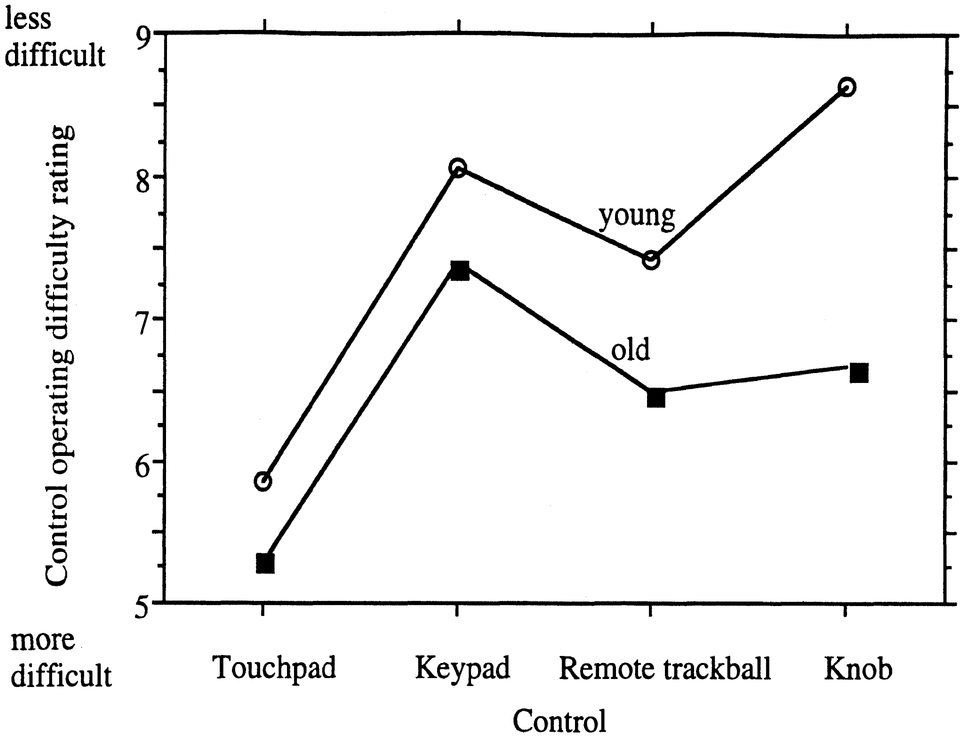


Figure 4.3.27. Differences due to age and control in ratings of operating difficulty

The model for ratings of difficulty of driving the simulator while operating controls was similar to that for the previous two ratings. Significant effects included age (p=.0017) and control (p=.0372). Consistent with previous measures, younger drivers rated the task less difficult than older drivers (6.7 versus 5.4), where 1=difficult to drive and 10=not difficult at all. Knobs were rated as least difficult to use while driving.

Combining these last three ratings, one can see that the pattern is consistent (Table 4.3.4). Overall, the knob was rated as the most usable control.

Table 4.3.4. Ratings of control usability in part 1

	Knob	Trackball	Keypad	Touchpad
Accessible	7.9	7.7	6.6	6.6
Easy to operate	7.6	6.9	7.7	5.5
Easy to drive	6.5	6.3	6.1	5.0
Mean	7.3	7.0	6.8	5.7

4.4 DISCUSSION

The first experiment established that the UMTRI driving simulator was a reasonable context for assessing the use of secondary controls. Secondary task time measures from both contexts were remarkably similar. However, while there were some differences in the absolute values of driving measures between the two contexts, relative differences were consistent.

The second experiment examined the role of practice in in-vehicle interface evaluations. That experiment showed significant improvements in performance with practice for almost all measures. However, there were few interactions with practice, suggesting that data from early in practice could be used to make decisions as to which of several alternative interfaces was easiest to use. As a consequence of these two experiments, the UMTRI driving simulator was used for an experiment concerning hierarchical menu system design. Given the results of the second experiment, there was only one test session.

From the results of the third experiment, the number keypad seemed to be the best control for secondary systems, with the knob being a second choice. Of the secondary task performance measures, the keypad always had the smallest interbutton RT and task completion time among the four controls, with times being about 1/3 less than those for knobs, the second choice. Furthermore, the Hick's Law analysis indicated that the gap between the keypad and other alternatives increased as the number of menu options increased.

The glance data also favor the keypad. The number keypad had the shortest number of glances to the IP and glances to the interface were shorter, on average, than for other alternatives. The second choice was the knob. Differences were on the order of several percent. A short IP glance duration was desired since drivers could spend more time looking at the forward scene. If the IP glance duration required by the secondary task was long, time pressure and uncertainty would build up and the drivers would be compelled to look forward again (Wierwille, 1993).

Prior to the experiment, the keypad was expected to require the greatest visual demand, since the drivers had to look at two displays (both the LCD and the control interface) to complete the secondary task. However, results from the experiment showed that the keypad actually required the smallest IP glance duration and task completion time. The small IP glance duration and task completion time might be due to the different selection procedures of the keypad. The keypad was an open-loop system and required direct manipulation. When a stimulus appeared on the display, drivers selected the item with one button push. This minimized the time needed to complete the task. On the other hand, the other three controls were closed-looped systems and they required multiple manipulations to complete one selection. For all of those controls, the drivers followed the cursor or the highlight constantly to make sure they were heading in the correct direction. Since multiple moves of the cursor or the highlight were usually required to make a selection, much longer IP glance durations and task completion times were needed.

As for the subjective ratings of selection difficulty, the keypad was preferred by the older age group. The younger age group voted it second. Younger drivers preferred the rotary knob, but no significant difference existed between their preferences for the rotary knob and the keypad.

For other ratings, preferences were mixed. The knob (mounted in this experiment near the right hand) was rated as most accessible, while the keypad and touchpad (located on the center console), were rated as least accessible. This problem could be easily solved by moving the keypad to a more accessible location such as the steering wheel, or higher up and at an angle on the center console to avoid awkward wrist angle. The accessibility rating results suggest that both control type and location need to be considered when selecting a control.

In terms of ease of operation, the keypad was rated as easiest to operate, with the knob being the second choice. For ease of use while driving, the knob was the first choice and the keypad was the third choice.

The final sets of measures of interest were for driving performance. In brief, there were no practical differences between controls with regard to driving performance.

Thus, these data suggest that the number pad was the preferred control with the knob being a second choice. The keypad led to significantly shorter task times than alternative controls and had lower visual demands in terms of fixation durations and number of fixations. This was due, in part, to the direct mapping between the number presented and the location of the numeric keys, a characteristic that may not be present in real products. Subjective ratings were mixed, sometimes favoring the keypad, sometimes favoring a knob. The subjective results emphasized the importance of context on the final selection of a control for a particular real product. For example, the knob was both accessible and easy to operate because the location was well chosen (behind the steering wheel on a stalk) and tactile feedback was provided (minimizing visual demands). The outcome might have been different had another location or a knob with poor tactile feedback been evaluated.

The remote trackball and touchpad did not do as well with regard to interbutton RT, IP glance duration, and subjective ratings. The two controls required complicated manipulations (hence the increase in interbutton RT), and they were hard to control in the simulator. Since both controls were positioning devices, the visual demand from the highly spatial requirements competed with driving, another spatial task. There were not enough spatial resources to share between both tasks. In addition, system instability could be a problem with the remote trackball and the touchpad. Either a high or a low gain in the controls would complicate control by drivers. In a situation where the drivers had to share their attention between driving and the secondary task already, increasing the secondary task demand was not wise.

As for the menu structures, the choice of the best menu structure is open to debate. For example, interbutton RTs were shorter for 2 x 6 menus (typically about 4 seconds) than for 4 x 3 (about 5 seconds) and 8 x 2 menus (about 7 seconds). Since only two choices were presented in each menu in the 2 x 6 menu, selections could be easily made. However, in terms of task completion times, there was almost no difference between 4 x 3 and 8 x 2 menus (both about 15 seconds), although both were considerably faster than 2 x 6 menus (about 25 seconds). For reasons of safety, the desire is to minimize both the total task time and the time required for each epoch of the in-vehicle task (when the driver is not attending to the road).

In terms of glance measures, glances to the 8 x 2 display (about 1.0 second) were longer than those to the 4 x 3 (under 0.9 seconds) and 2 x 6 (about 0.8 seconds) displays. The amount of information presented on each 2 x 6 menu was less than other menu structures, minimizing visual demand and subsequent mental processing. Hence, drivers did not need to spend too much time looking towards the IP at one time. This was advantageous since long IP glance durations can be unsafe. The most desirable condition would be to have the drivers spend as much time on the forward scene as possible. However, the 8 x 2 menus required far fewer glances (about 7) than 4 x 3 menu hierarchies (about 7.5) or 2 x 6 menu hierarchies (about 11). Hence, the total glance time away from the road was much greater for 2 x 6 menus, an undesired result. The tradeoff between the duration of individual glances and the number of glances on aggregate risk is unknown.

With regard to ratings of selection difficulty, the 2 x 6 menu was clearly easiest and the 8 x 2 menu was clearly most difficult. More overall measures of usability for menus should also have been considered. Additional, more detailed analysis of driving performance that examined the impact of menu structure was beyond the scope of this project.

These results do not support any definitive conclusions as to which menu structure is best across all of the measures. Depending on the results are viewed, there is support for either 2 x 6, 4 x 3, and 8 x 2 menu structures. This issue may be resolved by further analysis of the driving data.

In contrast, these data indicate that a keypad was best type of the control for menu selection in a moving vehicle, with a knob being a strong second choice. The recommendation for a particular context will vary with mounting location and other implementation details. Not explored, but having an important bearing on menu recommendations, are how shortcuts might be implemented to provide ready access to commonly used menu functions buried in a tree.

REFERENCES

Blaauw, G. J. (1980). Driving Experience and Task Demands in Simulator and Instrumented Car: A Validation Study, Progress Report II (Report Number. IZF 1980-9), Soesterberg, Netherlands: Institute for Perception TNO.

Blaauw, G. J. (1982). Driving Experience and Task Demands in Simulator and Instrumented Car: A Validation Study, Human Factors, 24(4), 473-486.

Buck, J. R. and Cheng, S. W. (1993). Instructions and Feedback Effects on Speed and Accuracy with Different Learning Curve Models, IIE Transactions, 25 (6), 122-131.

Card, S. K., English, W. K. and Burr, B. (1978). Evaluation of Mouse, Rate-Controlled Isometric Joystick, and Text Keys for Text Selection on a CRT, Ergonomics, 21(8), 601-613.

Card, S. K., Moran. T. P. and Newell, A. (1983). The Psychology of Human-Computer Interaction, Hillsdale, NJ: Lawrence Erlbaum Associates.

Green, P. (1979). Automobile Multifunction Stalk Controls: Literature, Hardware and Human Factors Review (Technical Report UM-HSRI-79-78), Ann Arbor, MI: The University of Michigan Highway Safety Research Institute.

Green, P. (1995a). Automotive Techniques (chapter), In Weimer, J., (ed.) Research Techniques in Human Engineering (2nd ed.), New York, NY: Prentice-Hall, 165-208.

Green, P. (1995b). Measures and Methods Used to Assess the Safety and Usability of Driver Information Systems (Technical Report FHWA-RD-94-088), Washington, D.C.: U.S. Department of Transportation, Federal Highway Administration.

Green, P., Lin, B., and Bagian, T. (1993). Driver Workload as a Function of Road Geometry: A Pilot Experiment (Technical Report UMTRI-93-39), Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Hada, H. (1994). Drivers' Visual Attention to In-Vehicle Displays: Effects of Display Location and Road Type (Technical Report UMTRI-94-9), Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Heuchert, S. (1995). Eyes Forward - an Ergonomic Solution to Driver Information Overload (Paper 9531183), The Eighth International Pacific Conference on Automotive Engineering (IPC-8), Tokyo, Japan: Society of Automotive Engineers of Japan, 169-174.

Hicks, T. G. and Wierwille, W. W. (1979). Comparison of Five Mental Workload Assessment Procedures in a Moving-Base Driving Simulator, Human Factors, 21(2), 129-143.

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 Advances in Industrial Ergonomics and Safety I, Bristol, PA: Taylor and Francis, 451-456.

Kantowitz, B. H. (1995). Simulator Evaluation of Heavy-Vehicle Driver Workload. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meetings, Santa Monica, CA: Human Factors and Ergonomics Society, 1107-1111.

Kiefer, R. J. and Gellatly, A. W. (1996). Quantifying the Consequences of the "Eyes-on-Road" Benefit Attributed to Head-Up Displays (SAE Paper 960946), Warrendale, PA: Society of Automotive Engineers.

Kurokawa, K. and Wierwille, W.W. (1990). Validation of a Driving Simulation Facility for Instrument Panel Task Performance, Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica, CA: Human Factors and Ergonomics Society, 1299-1303.

Landauer, T.K. and Nachbar, D.W. (1985). Selection From Alphabetic and Numeric Menu Trees Using A Touch Screen: Breadth, Depth, and Width, CHI'85 Proceedings, New York, NY: The Association of Computing Machinery, 73-78.

Li, C., Sumie, M., and Green, P. (1997). Repeatability of Secondary Task Workload Assessments between Driving Simulators in Two Countries (Technical Report 97-17), Ann Arbor, MI: The University of Michigan Transportation Research Institute (in progress).

MacAdam, C. C., Green, P., and Reed, M. P. (1993). An Overview of Current UMTRI Driving Simulators, UMTRI Research Review, 24(1), 1-8.

Miller, D.P. (1980). Factors Affecting Item Acquisition Performance in Hierarchical Systems: Depth vs. Breadth (unpublished Ph.D. dissertation), Columbus, OH: Ohio State University.

Miller, D.P. (1981). The Depth/Breadth Tradeoff in Hierarchical Computer Menus, Proceedings of the Human Factors Society 25th Annual Meeting, Santa Monica, CA: Human Factors and Ergonomics Society, 296-300.

Musseler, J. (1994). Using Predictors to Partition Menu Selection Times, Behavior and Information Technology, 13(6), 362-372.

Norman, K.L. (1991). The Psychology of Menu Selection, Norwood, NJ: Ablex.

Noy, Y. I. (1990). Attention and Performance While Driving With Auxiliary In-Vehicle Display (Report Number TP 10727 (E)), Ottawa, Ontario, Canada: Road Safety and Motor Vehicle Regulation Office, Transport Canada.

Prynne, K. (1995). Tactile Controls. Automotive Interiors International, Summer 1995, 33-36.

Reed, M. P. and Green, P. (1995). Validation of a Low-Cost Driving Simulator Using a Telephone Dialing Task (Technical Report UMTRI-95-19), Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Sivak, M. (1996). The Information that Drivers Use: Is It Indeed 90 Percent Visual? Perception, 25, 1081-1089.

Wickens, C. D. (1992). Engineering Psychology and Human Performance (2nd ed.), New York, NY: Harper Collins.

Wierwille, W. W. (1993). Visual and Manual Demands of In-Car Controls and Displays. In Peacock, B. and Karwowski, W. (eds.), Automotive Ergonomics, Bristol, PA: Taylor and Francis, 299-320.

Wierwille, W. W. and McFarlane, J. (1991). Overview of a Study on Direction-of Motion Stereotype Strengths for Automobile Controls (SAE Paper 910115), Warrendale, PA: Society of Automotive Engineers.

Zwahlen, H.T. (1993). Evaluation of Pushbutton Arrangements in Automobile. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, Santa Monica, CA: Human Factors and Ergonomics Society, 969-973.

APPENDIX A - CONSENT FORM

- Multiple Switch Study -
Participant Consent Form

We are working on new designs for multiple switches on instrument panel that will make them easier to use. Well designed car multiple switches can be used at a glance, so people can concentrate on driving. Responses from typical drivers such as you, will help identify the best multiple switch design.

You will be driving a simulated car in the laboratory at UMTRI. Your task will be to select the six numbers on a display using four different controls, such as a keypad, a trackball, a touchpad, and a rotary knob. The numbers will be instructed verbally.

The entire study will take about 2 hours to complete. You will be paid $35 for your participation. You should be aware that you can withdraw from the study at any time and for any reason. You will be paid regardless.

We will videotape the session with your permission. We will not release any identifying information, so your responses will remain confidential.

Thank you for your help with our study. If you have any questions, please do not hesitate to ask the experimenter before signing this form.

I have reviewed and understand the information presented above. My participation in this study is entirely voluntary.

Subject Name (PRINTED)	Date
Subject Signature	Witness (Experimenter)

it is OK to videotape me: yes no (circle one)

Investigator: Minoru Sumie 763-2485

APPENDIX B - BIOGRAPHICAL FORM

University of Michigan Transportation Research Institute
Human Factors Division
Multiple Switch Study Biography Form

Subject: ____________

Date: ____________

Name: ____

Male Female (circle one) Age: ____________

Occupation: ____________________

What kind of car do you drive the most?

year: __________ make: ______________ model: __________

Annual mileage: ______________

Have you ever driven the driving simulator at UMTRI? yes no

How susceptible are you to motion sickness?

never get motion sick	moderately don't get motion sick	neutral	moderately motion sick	frequently get motion sick

---

Titmus Vision: (Landolt Rings)

Vision Correction: Yes (Eye Glasses, Hard Contact Lens, Soft Contact Lens), No

1	2	3	4	5	6	7	8	9	10	11	12	13	14
T	R	R	L	T	B	L	R	L	B	R	B	T	R
20/200	20/100	20/70	20/50	20/40	20/35	20/30	20/25	20/22	20/20	20/18	20/17	20/15	20/13

APPENDIX C - DIFFICULTY RATING FORM

Experiment Difficulty Rating (1)

Name:________________

Driving Simulator ___35 mi/hr ___50 mi/hr

1. On the scale below, circle the numbers that best describe **how difficult it was to select numbers from the three menu displays.**

1	Back	2

Menu (1)

	12	
45	Back	23
	34	

Menu (2)

123	234	345
812	Back	456
789	678	567

Menu (3)

Menu (1):

	extremely difficult									*not at all difficult*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

Menu (2):

	extremely difficult									*not at all difficult*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

Menu (3):

	extremely difficult									*not at all difficult*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

Name:________________

2. On the scale below, circle the numbers that best describe **how accessible the controls were.**

	hardly accessible									*easy accessible*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

3. On the scale below, circle the number that best describe **how difficult it was to operate the controls.**

	extremely difficult									*not at all difficult*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

4. On the scale below, circle the numbers that best describe **how difficult it was to drive the simulator while operating the controls.**

	extremely difficult									*not at all difficult*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

Experiment Difficulty Rating (2)

Name:________________

1. On the scale below, circle the numbers that best describe **your preferences for the three menu displays.**

1	Back	2

Menu (1)

	12	
45	Back	23
	34	

Menu (2)

123	234	345
812	Back	456
789	678	567

Menu (3)

	not at all preferable				*average*					*extremely preferable*
Menu (1)	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Menu (2)	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Menu (3)	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

2. On the scale below, circle the numbers that best describe **your preferences for the four controls.**

	not at all preferable				*average*					*extremely preferable*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

3. On the scale below, circle the numbers that best describe **how safe you felt when driving the simulator and operating the controls.**

	not at all safe				*average*					*extremely safe*
Touchpad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Keypad	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Trackball	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Knob	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

Thank you for your cooperation!

APPENDIX D - DATA SUMMARY

Table 5.1. Data summary of simulator validity experiment (hand movement measures)

		Interbutton RT (s)	Task completion time (s)
Age	young	1.780	8.961
	old	2.209	11.503
Gender	male	1.960	10.045
	female	1.950	9.952
Context	on-road	1.946	9.956
	simulator	1.967	10.067
(On-road)	rural	2.053	10.631
	residential	1.969	10.326
	suburban	1.792	9.040
	expressway	1.976	9.878
(Simulator)	35 mi/hr	1.893	9.872
	50 mi/hr	1.954	9.770
	75 mi/hr	2.058	10.561
Mean		1.105	10.232

Table 5.2. Data summary of simulator validity experiment (eye fixation measures)

		IP glance duration (s)	Forward viewing time (s)	# of IP glances
Age	young	1.125	1.045	5.575
	old	0.966	1.227	6.793
Gender	male	1.003	1.176	6.272
	female	1.115	1.073	5.789
Context	on-road	1.100	1.021	6.106
	simulator	0.987	1.276	5.984
(On-road)	rural	1.074	1.037	6.450
	residential	1.109	1.087	5.952
	suburban	1.131	0.965	5.767
	expressway	1.087	1.002	6.267
(Simulator)	35 mi/hr	1.026	1.166	6.308
	50 mi/hr	1.014	1.312	5.860
	75 mi/hr	0.914	1.366	5.814
Mean		1.046	1.136	6.184

Table 5.3. Data summary of simulator validity experiment (driving performance)

		S.D. of steering angle (deg)	S.D. of lane position (%)	S.D. of throttle position (%)	S.D. of vehicle speed (mi/hr)
Age	young	10.988	10.4	10.092	2.716
	old	14.013	16.9	17.525	3.151
Gender	male	13.141	13.2	12.676	2.744
	female	11.366	12.9	13.593	3.038
Context	on-road	3.363	5.9	6.905	3.047
	simulator	21.374	20.5	19.581	2.738
(On-road)	rural	2.359	5.8	4.294	2.199
	residential	4.962	6.1	4.446	1.892
	suburban	3.910	5.8	5.298	4.238
	expressway	2.258	5.8	4.890	1.589
(Simulator)	35 mi/hr	11.825	10.3	16.619	0.357
	50 mi/hr	20.366	18.9	30.229	0.803
	75 mi/hr	32.562	33.0	31.649	2.013
Task	with task	10.442	13.5	13.616	2.802
	without task	13.734	12.7	12.752	2.972
(On-road)	with task	2.325	4.8	3.445	1.270
	without task	4.229	6.7	5.820	3.499
(Simulator)	with task	18.743	22.4	21.349	0.689
	without task	23.603	18.9	29.933	1.340
Mean		12.501	13.65	13.809	2.934

Table 5.4. Data summary of learning experiment (hand movement measures)

		Interbutton RT (s)	Task completion time (s)
Subject	subject A	1.607	8.057
	subject B	1.369	6.841
Condition	35 mph with task	1.467	7.348
	50 mph with task	1.507	7.537
Mean		1.488	7.449

Table 5.5. Data summary of learning experiment (driving performance)

		S.D. of steering angle (deg)	S.D. of lane position (%)	S.D. of throttle position (%)	S.D. of vehicle speed (mi/hr)
Subject	subject A	20.576	13.3	0.504	0.291
	subject B	19.761	10.4	0.582	0.268
Condition	35 mi/hr with task	15.971	10.6	0.472	0.309
	50 mi/hr with task	22.498	12.9	10.553	2.831
	50 mi/hr without task	22.116	11.9	10.055	2.878

Table 5.6. Data summary of compatibility experiment (hand movement measures)

		Interbutton RT (s)	Task completion time (s)
Age	young	4.158	16.117
	old	5.609	22.467
Gender	male	4.706	18.393
	female	5.042	20.001
Speed	35 mi/hr	4.832	18.884
	50 mi/hr	4.903	19.453
Control	keypad	3.760	14.756
	knob	4.948	19.794
	trackball	5.332	19.862
	touchpad	5.628	23.019
Menu	2 x 6	3.843	25.147
	4 x 3	5.370	17.207
	8 x 2	7.226	15.164
Mean		4.884	19.292

Table 5.7. Data summary of compatibility experiment (eye fixation measures)

		IP glance duration (s)	Forward-viewing time (s)	# of IP glances
Age	Young	0.808	1.059	8.088
	Old	1.069	1.288	9.183
Gender	Male	0.920	1.039	8.456
	Female	0.891	1.252	8.442
Speed	35 mi/hr	0.985	1.083	8.330
	50 mi/hr	0.830	1.197	8.572
Control	Keypad	0.779	1.111	7.676
	Knob	0.858	1.129	8.486
	Trackball	0.946	1.162	8.936
	Touchpad	1.021	1.159	8.709
Menu	2 x 6	0.848	1.146	11.187
	4 x 3	0.906	1.149	7.652
	8 x 2	1.013	1.129	6.529
Mean		0.939	1.174	8.636

Table 5.8. Data summary of compatibility experiment (driving performance)

		S.D. of steering angle (deg)	S.D. of lane position (%)	S.D. of throttle position (%)	S.D. of vehicle speed (mi/hr)
Age	young	20.881	25.5	1.916	1.359
	old	24.483	85.3	5.577	3.257
Gender	male	22.836	76.8	4.119	2.095
	female	22.326	28.7	3.146	2.444
Speed	35 mi/hr	19.855	66.5	2.376	1.885
	50 mi/hr	25.615	40.1	5.070	2.676
Control	keypad	22.423	95.4	3.103	2.138
	knob	22.640	45.2	4.315	2.447
	trackball	21.442	19.3	3.696	2.074
	touchpad	23.873	54.8	3.509	2.383
Mean		22.647	55.4	2.713	2.308

APPENDIX E - ANOVA TABLES

Table 5.9. ANOVA summary of task time measures

Source	df	Interbutton RT (s) F	p	Task completion time (s) F	p
Age (A)	1	131.335	0.0001	132.884	0.0001
Gender (G)	1	7.835	0.0052	7.068	0.0083
Subject (S) -nested	4	28.641	0.0001	19.578	0.0001
Context (C)	1				
A x G	1	19.912	0.0001	8.844	0.0032
A x C	1				
G x C	1				
S x C	5	3.933	0.0007		

(Probability values of less than 0.05 are shown)

Table 5.10. ANOVA summary of eye fixation measures

Source	df	IP glance duration (s) F	p	Forward-viewing time (s) F	p	Number of IP glances F	p
Age (A)	1	112.936	0.0001	30.593	0.0001	76.509	0.0001
Gender (G)	1	26.743	0.0001	31.540	0.0001	23.346	0.0001
Subject (S) -nested	4	38.851	0.0001	32.580	0.0001	16.072	0.0001
Context (C)	1	36.788	0.0001	61.647	0.0001		
A x G	1	60.005	0.0001	6.170	0.0131	7.235	0.0076
A x C	1	8.782	0.0031	6.147	0.0133		
G x C	1						
S x C	5	6.324	0.0001	6.297	0.0001	2.496	0.0228

(Probability values of less than 0.05 are shown)

Table 5.11. ANOVA summary of lateral control measures

		S.D. of steering angle		S.D. of lane position	
Source	df	F	p	F	p
Age (A)	1	28.931	0.0001	34.063	0.0001
Gender (G)	1				
Subject (S) -nested	4				
Context (C)	1	1171.488	0.0001	493.042	0.0001
Secondary task (T)	1	35.859	0.0001	7.522	0.0063
A x G	1				
A x C	1	4.510	0.0342	64.276	0.0001
A x T	1	6.855	0.0091		
G x C	1				
G x T	1				
S x C	1			2.977	0.0189
S x T	5				
C x T	3			11.250	0.0009

(Probability values of less than 0.05 are shown)

Table 5.12. ANOVA summary of longitudinal control measures

		S.D. of throttle position		S.D. of vehicle speed	
Source	df	F	p	F	p
Age (A)	1	7.691	0.0057	29.691	0.0001
Gender (G)	1	24.123	0.0001	8.308	0.0041
Subject (S) -nested	4	13.623	0.0001	7.635	0.0001
Context (C)	1	343.667	0.0001	328.804	0.0001
Secondary task (T)	1	30.721	0.0001	55.626	0.0001
A x G	1	5.011	0.0256	77.439	0.0001
A x C	1	12.917	0.0004	40.098	0.0001
A x T	1				
G x C	1	4.858	0.0279	7.371	0.0068
G x T	1			4.442	0.0355
S x C	1	14.301	0.0001	22.922	0.0001
S x T	5				
C x T	3	7.636	0.0059		

(Probability values of less than 0.05 are shown)

Table 5.13. ANOVA summary of interbutton RT (individual)

	On-Road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	99.471	0.0001	1	39.129	0.0001
Gender (G)	1			1	5.658	0.0177
Subject (S) -nested	4	19.259	0.0001	4	16.380	0.0001
Context (C)	3	9.044	0.0001	1		
A x G	1	21.748	0.0001	1	5.726	0.0170
A x C	3	7.988	0.0001	1		
G x C	3			1		
S x C	15	2.696	0.0002	10	2.082	0.0163

(Probability values of less than 0.05 are shown)

Table 5.14. ANOVA summary of task completion time (individual)

	On-Road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	77.059	0.0001	1	59.540	0.0001
Gender (G)	1			1	4.272	0.0412
Subject (S) -nested	4	14.671	0.0001	4	9.363	0.0001
Context (C)	3	5.570	0.0012	1		
A x G	1	10.336	0.0016	1		
A x C	3	3.332	0.0214	1		
G x C	3			1		
S x C	15			10		

(Probability values of less than 0.05 are shown)

Table 5.15. ANOVA summary of IP glance duration (individual)

	On-Road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	105.947	0.0001	1	26.375	0.0001
Gender (G)	1	21.370		1	9.270	0.0024
Subject (S) -nested	4	14.373	0.0001	4	32.229	0.0001
Context (C)	3	3.419	0.0169	1	4.964	0.0072
A x G	1	30.341	0.0001	1	29.928	0.0001
A x C	3			1		
G x C	3	3.476	0.0156	1		
S x C	15	2.359	0.0011	10		

(Probability values of less than 0.05 are shown)

Table 5.16. ANOVA summary of forward-viewing time (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	9.806	0.0018	1	26.153	0.0001
Gender (G)	1	39.422	0.0001	1	5.972	0.0148
Subject (S) -nested	4	38.568	0.0001	4	5.603	0.0001
Context (C)	3			1	3.439	0.0327
A x G	1	21.854	0.0001	1		
A x C	3			1		
G x C	3	3.120	0.0254	1		
S x C	15	3.054	0.0001	10		

(Probability values of less than 0.05 are shown)

Table 5.17. ANOVA summary of number of IP glances (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	95.501	0.0001	1	17.186	0.0001
Gender (G)	1	11.987	0.0007	1	13.878	0.0003
Subject (S) -nested	4	22.072	0.0001	4	3.598	0.0050
Context (C)	3	3.357	0.0209	1		
A x G	1			1		
A x C	3	3.668	0.0140	1		
G x C	3			1		
S x C	15			10		

(Probability values of less than 0.05 are shown)

Table 5.18. ANOVA summary of S.D. of steering angle (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1	4.456	0.0357	1	46.377	0.0001
Gender (G)	1			1		
Subject (S) -nested	4			4	2.333	0.0430
Context (C)	3	25.655	0.0001	2	80.792	0.0001
Secondary task (T)	1	39.106	0.0001	1	24.732	0.0001
A x G	1			1		
A x C	3			2		
A x T	1			1	12.523	0.0005
G x C	3			2		
G x T	1			1		
S x C	15			10		
S x T	5			3		
C x T	3			1		

Table 5.19. ANOVA summary of S.D. of standardized lane position (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1			1	103.564	0.0001
Gender (G)	1			1		
Subject (S) -nested	4	3.484	0.0164	4		
Context (C)	3			2	96.670	0.0001
Secondary task (T)	1	52.037	0.0001	1		
A x G	1			1		
A x C	3			2		
A x T	1			1		
G x C	3			2		
G x T	1			1		
S x C	15			10		
S x T	5			3	2.630	0.0173
C x T	3			1		

(Probability values less than 0.05 are shown)

Table 5.20. ANOVA summary of S.D. of throttle position (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1			1	24.325	0.0001
Gender (G)	1			1	27.508	0.0001
Subject (S) -nested	4	9.902	0.0001	4	11.946	0.0001
Context (C)	3	3.287	0.0213	2	4.237	0.0156
Secondary task (T)	1	43.955	0.0001	1		
A x G	1	13.318	0.0003	1	28.949	0.0001
A x C	3			2		
A x T	1			1		
G x C	3			2		
G x T	1			1	4.581	0.0333
S x C	15			10	4.085	0.0001
S x T	5			3		
C x T	3			1		

(Probability values less than 0.05 are shown)

Table 5.21. ANOVA summary of S.D. of vehicle speed (individual)

	On-road			Simulator		
Source	df	F	p	df	F	p
Age (A)	1			1	58.051	0.0001
Gender (G)	1			1	12.098	0.0006
Subject (S) -nested	4			4	13.913	0.0001
Context (C)	3	20.247	0.0001	2	29.891	0.0001
Secondary task (T)	1	126.885	0.0001	1	15.559	0.0001
A x G	1	6.437	0.0118	1	99.510	0.0001
A x C	3			2		
A x T	1			1		
G x C	3			2	18.066	0.0001
G x T	1			1	6.201	0.0135
S x C	15			10	2.309	0.0084
S x T	5			3		
C x T	3	3.142	0.0258	1		

(Probability values of less than 0.05 are shown)

Table 5.22. ANOVA summary of secondary task performance

		Interbutton RT		Task completion time	
Source	df	F	p	F	p
Subject (S)	1	117.835	0.0001	504.473	0.0001
Condition (C)	1	4.635	0.0377		
Day (D)	19	35.803	0.0001	35.245	0.0001
S x C	1				
S x D	18	3.835	0.0001	4.168	0.0001
C x D	19				

(Probability values of less than 0.05 are shown)

Table 5.23. ANOVA summary of lateral control measures

		S.D. of steering angle		S.D. of lane position	
Source	df	F	p	F	p
Subject (S)	1	82.899	0.0001	60.591	0.0001
Condition (C)	2	2561.947	0.0001	15.021	0.0001
Day (D)	19	3.107	0.0019	3.078	0.0021
S x C	2	19.713	0.0001	28.201	0.0001
S x D	18				
C x D	38	2.162	0.0124		

(Probability values of less than 0.05 are shown)

Table 5.24. ANOVA summary of longitudinal control measures

Source	df	S.D. of throttle position F	S.D. of throttle position p	S.D. of vehicle speed F	S.D. of vehicle speed p
Subject (S)	1				
Condition (C)	2				
Day (D)	19	3.201	0.0015		
S x C	2				
S x D	18	2.087	0.0316	2.513	0.0101
C x D	38				

(Probability values of less than 0.05 are shown)

Table 5.25. ANOVA summary of task time measures

Source	df	Interbutton RT F	Interbutton RT p	Task completion time F	Task completion time p
Age (A)	1	85.254	0.0001	170.228	0.0001
Gender (G)	1			24.262	0.0001
Subject (S) -nested	12	5.330	0.0001	15.414	0.0001
Speed (Sp)	1	12.136	0.0005	22.749	0.0001
Control (C)	3	32515	0.0001	43.007	0.0001
Menu (M)	2	190.646	0.0001	140.097	0.0001
A x G	1				
A x Sp	1				
A x C	3			2.769	0.0411
A x M	2				
G x Sp	1				
G x C	3	4.104	0.0065		
G x M	2				
S x Sp	13			3.442	0.0001
S x C	39			1.937	0.0013
S x M	26			1.539	0.0443
Sp x C	3				
Sp x M	2				
C x M	6	3.991	0.0006	3.070	0.0058

(Probability values of less than 0.05 are shown)

Table 5.26. ANOVA summary of eye fixation measures

Source	df	IP glance duration F	IP glance duration p	Forward-viewing time F	Forward-viewing time p	Number of IP glances F	Number of IP glances p
Age (A)	1	15.482	0.0001	87.076	0.0001	20.526	0.0001
Gender (G)	1	4.108	0.0427	61.736	0.0001		
Subject (S) -nested	12	24.291	0.0001	23.612	0.0001	9.585	0.0001
Speed (Sp)	1	10.832	0.0010	5.842	0.0157		
Control (C)	3	34.622	0.0001			4.446	0.0043
Menu (M)	2	18.580	0.0001	5.722	0.0033	136.784	0.0001
A x G	1					4.617	0.0321
A x Sp	1	7.777	0.0053				
A x C	3	4.484	0.0038				
A x M	2						
G x Sp	1	6.887	0.0087				
G x C	3			3.844	0.0092		
G x M	2						
S x Sp	13	4.288	0.0001	2.244	0.0170		
S x C	39	4.371	0.0001	2.077	0.0005	1.748	0.0094
S x M	26						
Sp x C	3	6.365	0.0003				
Sp x M	2						
C x M	6	2.706	0.0127	2.107	0.0493		

(Probability values of less than 0.05 are shown)

Table 5.27. ANOVA summary of lateral control measures

Source	df	S.D. of steering angle F	S.D. of steering angle p	S.D. of lane position F	S.D. of lane position p
Age (A)	1	112.771	0.0001	21.113	0.0020
Gender (G)	1				
Subject (S) -nested	12	22.939	0.0001	4.990	0.0001
Speed (Sp)	1	293.683	0.0001		
Control (T)	3	5.700	0.0024		
A x G	1	10.126	0.0028		
A x Sp	1				
A x C	3				
G x Sp	1				
G x C	3				
S x Sp	12	2.591	0.0120		
S x C	39	2.056	0.0128		
Sp x C	3				

(Probability values of less than 0.05 are shown)

Table 5.28. ANOVA summary of longitudinal control measures

		S.D. of throttle position		S.D. of vehicle speed	
Source	df	F	p	F	p
Age (A)	1	30.839	0.0001	33.745	0.0001
Gender (G)	1			14.531	0.0005
Subject (S) -nested	12	16.382	0.0001	12.060	0.0001
Speed (Sp)	1	67.888	0.0001	26.988	0.0001
Control (T)	3			2.903	0.0465
A x G	1	11.998	0.0013		
A x Sp	1				
A x C	3				
G x Sp	1	14.440	0.0005	6.808	0.0127
G x C	3				
S x Sp	12				
S x C	39				
Sp x C	3				

(Probability values of less than 0.05 are shown)

Table 5.29. ANOVA summary of selection difficulty ratings

		Subjective ratings	
Source	df	F	p
Age (A)	1	103.626	0.0001
Gender (G)	1	12.932	0.0004
Subject (S) -nested	12	15.874	0.0001
Speed (Sp)	1	11.351	0.0008
Control (C)	3	11.663	0.0001
Menu (M)	2	4.654	0.0103
A x G	1	13.322	0.0003
A x Sp	1		
A x C	3	4.144	0.0067
A x M	2		
G x Sp	1	7.187	0.0067
G x C	3		
G x M	2		
Sp x C	3		
Sp x M	2		
C x M	6		

(Probability values of less than 0.05 are shown)