Driver interaction with different in-vehicle navigation system input devices: a validation study comparing simulation and on-road testing

Ying Wang, Vincent Lammers, Bryan Reimer, Bruce Mehler, Lisa A. D'Ambrosio & Joseph F. Coughlin

MIT AgeLab — Ideas + Technology for Quality Living

New England University Transportation Center

ABSTRACT

In-vehicle information systems are rapidly moving from novelties to becoming a common feature of both consumer and fleet vehicles. Driving simulation has been offered as a method of evaluating the functional design, ergonomics, cognitive demands and safety of in-vehicle information system interfaces. Simulation is appealing since it provides a safe environment for conducting such testing. However, relatively little research has addressed the extent to which interactions with in-vehicle information systems in the simulator provides an accurate model of behavior in the real world. In this study, we assess the validity of driving simulation for assessing the demands of various methods of interacting with in-vehicle systems. A surrogate destination entry task was developed based upon a typical interaction with modern text messaging system. In our surrogate task, we required drivers to enter the first two letters of a road name, city, and state with the system automatically completing the remainder of the entry. Three interfaces were evaluated: two paralleling methodologies used by various automobile manufacturers (touch screen and distributed rotational controller) and a third based upon a cellular phone keypad. Results from an on-road field study were compared with those obtained from a stationary single screen driving simulator to assess the validity of measures of visual attention, driving performance, and task performance. Two independent sets of participants took part in the study, 31 in the driving simulator and 33 in the field study. Participants were between the ages of 22 and 28. During the destination entry task, most visual attention measures and destination entry task performance indices, such as glance frequency, total glance duration, and mean task response time, mapped almost identically from simulation to field. The driving performance measures, including standard deviation of velocity and lateral offset, appear somewhat less consistent for discriminating between devices; however, there were no statistically significant differences between the two experimental environments except for a modest magnitude difference in the standard deviation of lane position in simulation vs. on-road. In summary, we found that visual attention, task performance and impact on driving performance mapped well between simulation and the field. In conclusion, fixed based driving simulation appears to be a very effective method of evaluating task interaction, performance, and modeling of field behavior of in-vehicle device interfaces.

BACKGROUND

- The development of in-vehicle infotainment systems provides drivers with entertainment, communication capabilities and up-to-date navigation information. While there has been a great deal of positive interest and rapid adoption, there as been a concurrent increase in the level of concern regarding the safety of these devices. This safety concern applies to both potential driver distraction and total cognitive workload.
- Simulated driving provides a safe, replicable, and controllable experimental environment for evaluating the interfaces for in-vehicle information systems.
- While the validity of simulated driving in reproducing on-road driving behavior has been addressed in a number of studies, relatively little research has assessed the validity of driver interaction with in-vehicle information systems in a simulated environment.
- In this study, a surrogate destination entry task was used to test the validity of simulated driving in assessing the demands of various methods of interacting with in-vehicle systems.

METHODS

Participants

- 62 subjects between the ages of 22-28 were enrolled in the two independent samples. One simulator subject and 3 field subjects were dropped from analysis due to simulator sickness and recording issues respectively, resulting in 28 usable cases in the field sample and 30 in the simulator sample. Gender representation was relatively balanced in each group.
- Subjects were required to have held a valid driver's license for at least 3 years, a clean driving record for the previous year, and no prior experience in a driving simulation study. Subjects also were required to have no neurological problems, be free of a mental disorder or a major medical condition such as cancer, not be on medication for hypertension, have taken no medication that causes drowsiness, and not require prescription glasses to drive.

Fig 4. Driver's view of S-IVIS touch screen and keypad entry devices

Apparatus

1. Test vehicle: MIT AgeLab "Aware Car" – a 2003 series Volvo XC 90 SUV instrumented vehicle with eye tracker, CAN (Controller Area Network), lane tracker and six cameras.
2. Simulator: MIT AgeLab driving simulator "Miss Daisy" - a fixed-base simulator comprised of a full cab 2001 Volkswagen Beetle running STISIM Build 2.10.03 at 30Hz.
3. Eye Tracking System: Seeing Machines FaceLab 4.2, with cameras mounted on both the test vehicle and the simulator.
4. Surrogate in-vehicle information systems (S-IVIS): three integrated input devices consisting of a XENARC 800-TSV touch screen, a USB number block keypad, and a PC-based Griffin Technology USB scrolling wheel (iDrive).

Destination Entry Task (DET)

- Each destination entry consisted of a state, a city and a street.
- Entry of only the first two letters of each word was required.
- Three destination entry tasks were carried out with each device.

Fig. 1. MIT AgeLab Aware Car

Fig. 2. MIT Agelab Simulator – Miss Daisy

Fig. 3. Eye tracking system mounted on the dashboard

Experimental Design

1. Independent variables
   - Experimental site (2 levels): Field vs. Simulator (between subject comparison).
   - Input device (3 levels): Touch Screen (TS) vs. Keypad (KP) vs. Scrolling Wheel (SW) (within subject comparison).

2. Dependent variables
   - Visual attention: Glance Frequency, Total Glance Duration
   - Driving performance: Standard Deviation of Longitudinal Velocity, Standard Deviation of Lane Position
   - DET performance: Mean Task Duration, Mean Task Response Time

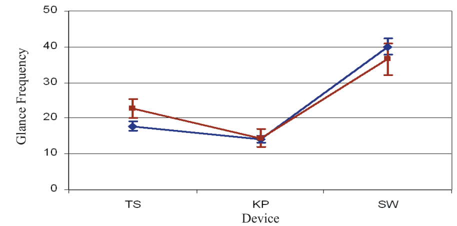
Data Analysis

General Linear Model with Repeated Measures (SPSS 11.5) with $p < .05$ used for establishing significance.
A Greenhouse-Geisser correction was applied in cases where the assumption of sphericity was not met.

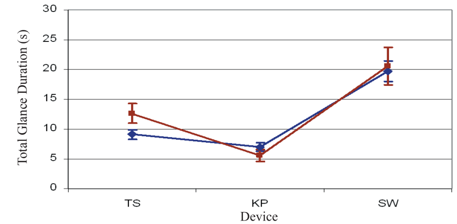
- Relative validity: defined as no significant interaction between site and device.
- Absolute validity: defined as no significant interaction between site and device, and no significant main effect for site.

RESULTS

Visual Attention

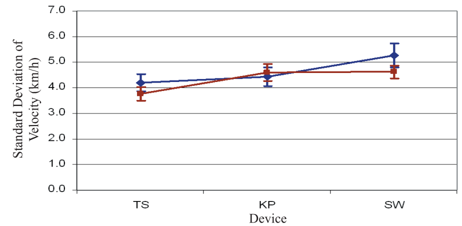


Interaction (site*device)	F (1.66, 79.48)=2.93	P=0.069
Main effect (site)	F (1, 48)=0.04	P=0.836

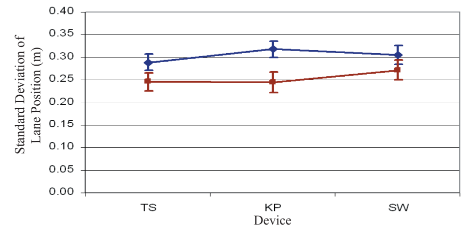


Interaction (site*device)	F (1.32, 64.58)=2.16	P=0.140
Main effect (site)	F (1, 49)=0.28	P=0.597

Driving Perfomance

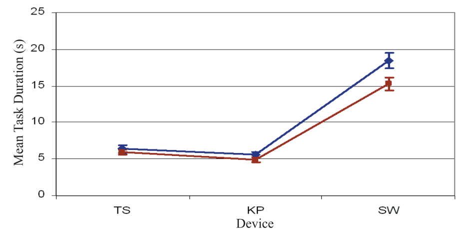


Interaction (site*device)	F (2, 110)=1.20	P=0.304
Main effect (site)	F (1, 55)=0.58	P=0.448

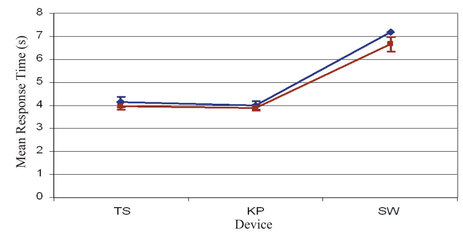


Interaction (site*device)	F (2, 94)=0.87	P=0.424
Main effect (site)	F (1, 47)=4.80	P=0.033

DET Task Performance



Interaction (site*device)	F (1.38, 63.24)=2.81	P=0.086
Main effect (site)	F (1, 46)=5.99	P=0.018



Interaction (site*device)	F (1.63, 78.07)=0.89	P=0.394
Main effect (site)	F (1, 48)=0.84	P=0.365

■ Simulator ■ Field

Analysis of Validity

	Measurements	Relative Validity	Absolute Validity
Visual Attention	Glance Frequency	Y	Y
	Total Glance Duration	Y	Y
Driving Performance	Standard Deviation of Longitudinal Velocity	Y	Y
	Standard Deviation of Lane Position	Y	N
DET Task Performance	Mean DET Task Duration	Y	N
	Mean DET Task Response Time	Y	Y

CONCLUSION

In conclusion, fixed based driving simulation is a safe method of assessing task interaction and performance that provides valid estimates of on-road behavior for the type of in-vehicle interface interaction examined in this study. Visual attention and task response time measures derived from the simulator appear particularly promising for modeling field behavior.

- Both relative and absolute validity hold for the visual attention measures.
- Both relative and absolute validity hold for the standard deviation of velocity measure of driving performance; relative validity is evident for standard deviation of lane position.
- Both relative and absolute validity hold for the mean destination entry task response time; relative validity is evident for the mean destination entry task duration while absolute validity is impacted due to slight but consistently higher task durations in the simulator. It could be argued that for purposes of comparing devices, a practical, if not absolute validity, was established for task duration.