

Comparing Three Novel Multimodal Touch Interfaces for Infotainment Menus

Richard Swette

Georgia Institute of Technology
654 Cherry Street
Atlanta GA 30332
+1.954.263.0201
rswette3@gatech.edu

Keenan R May

Georgia Institute of Technology
654 Cherry Street
Atlanta GA 30332
+1.832.247.4415
kmay@gatech.edu

Thomas M Gable

Georgia Institute of Technology
654 Cherry Street
Atlanta GA 30332
+1.404.894.2680
thomas.gable@gatech.edu

Bruce N Walker

Georgia Institute of Technology
654 Cherry Street
Atlanta GA 30332
+1.404.894.8265
bruce.walker@psych.gatech.edu

ABSTRACT

Three novel interfaces for navigating a hierarchical menu while driving were experimentally evaluated. Prototypes utilized redundant visual and auditory feedback (multimodal), and were compared to a conventional direct touch interface. All three multimodal prototypes employed an external touchpad separate from the infotainment display in order to afford simple eyes-free gesturing. Participants performed a basic driving task while concurrently using these prototypes to perform menu selections. Mean lateral lane deviation, eye movements, secondary task speed, and self-reported workload were assessed for each condition. Of all conditions, swiping the touchpad to move one-by-one through menu items yielded significantly smaller lane deviations than direct touch. In addition, in the serial swipe condition, the same time spent looking at the prototype was distributed over a longer interaction time. The remaining multimodal conditions allowed users to feel around a pie or list menu to find touchpad zones corresponding to menu items, allowing for either exploratory browsing or shortcuts. This approach, called *GRUV*, was ineffective compared to serial swiping and direct touch, possibly due to its uninterruptable interaction pattern and overall novelty. The proposed explanation for the performance benefits of the serial swiping condition was that it afforded flexible sub tasking and incremental progress, in addition to providing multimodal output.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: User Interfaces- Auditory (non-speech) feedback, Input devices and strategies, Interaction Styles, Voice I/O; H.1.2 [Models and Principles]: User/Machine Systems- Human factors.

General Terms

Performance, Design, Human Factors

Keywords

Multimodal interfaces. Distracted driving. Eyes-Free interaction. Sub-tasking. Menu Navigation.

1. INTRODUCTION

Navigating a hierarchical menu structure within in-vehicle technologies has become a common component of the increasingly multi-natured task of operating a motor vehicle. In-car activities

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
AutomotiveUI '13, October 28 - 30 2013, Eindhoven, Netherlands
Copyright 2013 ACM 978-1-4503-2478-6/13/10...\$15.00.
<http://dx.doi.org/10.1145/2516540.2516559>

now include interacting with navigation systems, entertainment systems, and information systems [1]. The increasing number of functions in-vehicle technologies (IVTs) must support has led to a reduction in the feasibility of 1:1 mapped tactile systems [2], and a shift toward generalized input devices like click knobs and direct touchscreens which can be used to navigate through menus of functions and commands.

The goal of this research was to further understand how multimodal feedback and streamlined gestural input can be used to allow drivers to perform menu navigation while minimizing the impact on driving and on their related ability to visually attend to the driving task. Despite its known dangers, many drivers do attempt to multitask, either using portable technologies such as smartphones or through systems designed specifically for the in-vehicle environment. Disabling IVTs while in motion is one possible solution, yet still presents the problem of drivers turning to unconnected smartphones, the use of which is both common and detrimental to driving performance [3].

Research into multitasking while driving has generally found that driving performance increases in the presence of feedback across multiple modalities (for review, see [4]). This avenue of investigation is becoming increasingly relevant as IVTs are built around touchscreens, which can lead to high visual demand. In addition to considering multimodal output, work has also been done to rethink IVT interactions using methods like as speech [5], handwriting recognition [6] and gestural interactions [7]. While many of these allow direct access to commands and functions, in some contexts users will need to traverse menus in the traditional manner. As such, the focus of the present work is on how to approach sequential menu navigation¹ with the novel use of touch devices and auditory cues in order to minimize distraction.

2. RELATED WORK

2.1 Multimodal Output in IVTs

The benefits of multimodality are attributed to the presence of parallel processing capacities or ‘multiple resources’ [8] where performance increases when information processing demands are distributed among separate resource pools. While broad measures of visual attention do not necessarily equate with situational awareness [9], visual inattention has been linked to increased reaction time to hazards [10] and the likelihood of accidents [11,12,13].

The general principle behind multimodality is that each modality stream is at the very least perceived and perhaps processed in parallel with the others, creating an information-rich environment

1. ‘Menu’ here refers to unsorted 2-12 item sets of selectable items, in contrast with ‘list’ which refers to long, sortable sets of items.

[8]. There are three general ways in which multimodality has been utilized in IVTs.

The first is to provide minimally informative multimodal augmentation to visual selection tasks. Pitts et al [14], Ranney et al. [15] and Lee and Spence [16] found benefits to driving performance for interfaces that added auditory confirmation to a visual task.

The second approach is to evaluate auditory-only displays. In some [17,18,19] but not all [20,21] studies that compared auditory-only interfaces to visual interfaces, there was an overall observed benefit to driving performance associated with auditory display, usually accompanied by slower secondary task times. Doubts remain as to whether auditory-only interfaces can work well inside complex menu structures, both because users may lose their bearings and because of potential increases to overall cognitive load, despite reduction in visual distraction [21,13]. Several studies indicate that advantages of auditory displays increase when driving task difficulty increases [17,18] and when secondary task difficulty is high [19,21].

This interaction effect fits with Jannsen and Brumby [23], who found that drivers tend to be adept at smoothly allocating attention between primary and secondary tasks as necessary. This behavior is supported by the third approach to multimodality—the use of ‘orthogonal’ multimodal displays. These are displays that can be used with either visuals or audio only, depending on driver preference and other factors. Much research into orthogonal multimodality in IVTs concludes that drivers can choose which modality output they will use to attend to the secondary task. Several studies found a driving performance advantage to orthogonal multimodal displays over audio-only, visual-only or both [17,21,22,24]. Others found no difference between the best of purely auditory or purely visual displays and multimodal displays [25,21,26]. Helleberg and Wickens [27] explain these findings as being due to drivers being unable to use the additional information in the way it was intended, with it instead becoming noise.

2.2 Alternate IVT Interactions

One proposed difficulty with auditory displays is that they can still sometimes lead to high cognitive workload despite their eyes-free operation, especially when used for search tasks instead of simple adjustments [13]. Recarte and Nunes [28] found that less visual scanning took place when participants performed an auditory secondary task. In light of this, interaction methods that would minimize overall load were considered, drawing both from interfaces for the visually impaired and novel IVT interfaces.

Direct touch systems have mass appeal and are expected to become more ubiquitous in vehicles over the next decade [29]. Yet, for a variety of reasons, touchscreens pose a danger to drivers. Green [30] advocates a search for input devices better suited for use while driving. A central pillar of IVT design is the need for systems that promote and allow sub tasking [30, 31, 32]. Direct touch systems do not require a discrete asynchrony between visual and motor movement, and the user must complete all stages of a selection task in one goal-directed movement subject to the error rate and attention demands of Fitts’s law [33]. While direct touch systems may be efficient overall, their design affords dangerously prolonged glances [34]. Rydström et al. [35] found that a serial rotary controller was preferable to direct touch in terms of dwell time during continuous adjustment tasks, but found no driving performance differences. Ecker et al. [36] tested a radial marking interface designed to minimize glancing and fine movement, observing shorter task times over direct touch but, again, no

difference in driving performance. Bach et al. [7] tested a gestural touchpad against direct touch and tactile buttons and observed shorter secondary task dwell time for the gestural condition, but no driving differences. Weinberg et al. [37] constructed but did not test a gestural navigation system emphasizing slow interaction speed and discrete steps. Modern multitouch trackpads can recognize various combinations of single or multi finger gestures, supporting gestural interactions as well as scrolling, marking or tracing ones. It also can be arbitrarily divided into touch regions in the manner of *Earpod* [38,39] or *Sliderule* [40]. Some of these modes can exist simultaneously. For example, absolute mapped touch zones could be used for menu traversal while gestures could manipulate content at menu endpoints executing functions like play or pause in a manner akin to Bach et al. [7].

Work has also been done to merge the efficiency of absolute-mapped interfaces like direct touch with auditory output. Both *Earpod* [38,39], a multimodal pie-shaped marking menu, and *Sliderule* [40] (an audio-only interface with targets arranged as horizontal bands) employ an interaction pattern in which the user explores with one finger around a touch device, listening to readouts about the current zone. These interactions allow both exploration and shortcutting based on memory of menu content. If the user is exploring, such interfaces allow for a continuous, user-controlled loop of input and output, as opposed to the punctuated, discrete sequence found in many auditory interfaces. The multimodal *Earpod* also allows for cross-modality shortcutting—users can glance briefly to estimate a target’s location, then move and correct using auditory readouts. Zhao et al [40] evaluated *Earpod* in the vehicle, and found pure auditory or orthogonal multimodal versions were less detrimental to driving performance than a visual-only condition.

3. DEVICE DESIGN

In an effort to investigate the potential for multimodal touchpad interfaces to reduce driver distraction, three such interfaces were designed—two absolute mapped, zone-based menus, collectively called *GRUV*, and a location-agnostic gestural interface, referred to as ‘serial swipe’ [Fig 1]. Two versions of *GRUV* were produced—one with ‘pie slice’ zones referred to as circle *GRUV* [Fig.2] and another with horizontal bar zones referred to as vertical *GRUV* [Fig.3]. In addition, a direct touch condition similar to existing commercial systems was created [Fig.4].

3.1.1 Visual Display

All designs leveraged the same set of pictorial icons, set against a white background. For the direct touch device, menus consisted of a set of 2-9 buttons, each 160x130px, distributed in a grid with 40px between each button. A back button, 100x200px, was located at the top left of the screen. Visual design of vertical *GRUV* and serial swiping was identical. The entirety of the screen was divided into a number of bands (the main menu was thus 9 bands, each 87 pixels high). Icons were reduced per the demands of the vertical menu. For the circular *GRUV*, menu items were distributed as equal slices of a pie. Selected items were highlighted in blue.

3.1.2 Interactions

In the direct touch condition, users single-tapped targets to select them, and pressed a back target to move back. For the remaining designs, there were two separate kinds of interactions—moving the selector and making a selection or returning to a previous menu.



Figure 1: serial swipe



Figure 2: circle *GRUV*



Figure 3: vertical *GRUV*



Figure 4: direct touch

For the *GRUV* prototypes, an item became selected when the user's finger entered the section of the touchpad corresponding to that item. If the user lifted their finger, this item remained selected, but nothing further occurred. Subsequently, placing a single finger on the device would make the system 'jump' the selection to the menu item corresponding to the new finger location. For the serial swipe prototype, the selector was moved up and down via two-finger swipes up or down. Users had to make one swipe per selection change, and could not smoothly scroll through items. These discrete steps were required in order to disallow free movement that might lead to the use of visuals to assess progress toward the goal, which can lead to driving performance degradation [37,21]. There was no limit on selector movement speed, but an error tone would play if the user tried to swipe past the top or bottom of the list. The choice to not allow wraparound was driven by pilot testers reporting a loss of bearing when using the system with only auditory output.

To make a selection or go back, two finger swipes to the left (back) and right (select) were used. Selection gestures could be performed anywhere on the trackpad. Other alternatives were experimented with, including one or two finger taps, press-and-tap, single and double taps. All of these proved either awkward or error prone in pilot testing, particularly in the two *GRUV* conditions. An error tone would play when participants tried to go back when at the home menu. For all three multimodal designs, the previous selector position was remembered when going forward or back.

During pilot testing, participants were sometimes confused as to where their fingers were on the trackpad. To remedy this, circular and rectangular acrylic cutaways approximately 1mm deep were overlaid over the trackpad, allowing participants to quickly find the interaction space, and feel around its edge instead of being lost in its middle. Some pilot participants also used the edges to quickly go to the first or last item in a vertical *GRUV* menu. These overlays also provided an interaction area for circle *GRUV*, small enough to discourage ulnar deviation and large enough to reduce the need for fine motor movement.

3.1.3 Auditory Display

When moving within a menu, auditory cues were used to provide information of the current selection as well as a subtle sense of context and spatial location within the menu without glancing. To indicate the current selection, accelerated (to about 175%) and interruptible text to speech [42] output was used to read out each menu item. Before each item was read out, a very short (15ms) 440hz tone was given followed by a 15ms gap. This tone was pitch shifted based off of menu position- as users traversed the menu starting from the top or from 90 degrees, the tone would rise in pitch.

To provide information about the content of a new menu, similar cues were used. A full or partial chromatic scale of tones, starting with 440hz, was played over 800ms. This 'overview' scale was designed to provide a sense of the size and scope of the menu, similar to the auditory scrollbars described by Yalla and Walker [41]. It was accompanied by a 'whoosh' sound, but not a readout of the new menu name, since this was redundant with the name spoken at the time of selector change and led to annoyance in pilot testing. When going back, the title of the menu was played at a lesser amplitude and higher pitch, and the 'whoosh' sound was played in reverse. The addition of these initial auditory outputs synced with a 400ms period after moving between menus when users were restricted from moving the selector.

4. EXPERIMENT

In an effort to evaluate the level of distraction posed by each of the four device prototypes while performing the driving task, this study employed a medium-fidelity simulator and simulated IVTs while measuring driving performance seen through mean lane deviation during multitasking phases, secondary task menu navigation performance, eye-tracking, subjective cognitive workload and subjective preferences.

4.1 Method

4.1.1 Participants

Participants consisted of 19 drivers (13 male, 7 female) who had normal or corrected to normal vision and hearing. All had a valid driver's license, were undergraduate students from a technical university in the southeastern US, aged 18-24. They were compensated with one credit hour, for one hour of their time.

4.1.2 Materials and Apparatus

For the driving task, the Lane Change Task (LCT) [43] was employed. Participants were tasked with changing lanes in accordance with cue signs spaced 150 meters apart while fully depressing the gas pedal, thus maintaining a steady 60 mph. Input was handled through a Logitech Driving Force GT racing wheel and floor pedal. Simulator output was displayed on a 40" Samsung UN40C6300 TV. The height of the steering wheel was approximately 7 inches from the desk, which was 42 inches high. The base of the car used was approximately 21" high and 12" from the edge of the desk, with the wheel protruding 6" from the desk [Fig.5].

An Apple Magic Trackpad and a small touchscreen head-unit were used to simulate the secondary task. The track pad was positioned on an armrest 38 inches high, and approximately 12 inches to the right and 16 inches ahead of the participant's right shoulder (inside their reach envelope). The secondary display was 800x600px and 6x4", and was positioned 5 inches behind and to the right of the base of steering wheel- just outside the reach envelope (without leaning forward) of most participants. Prototype software was constructed in Adobe Flash, and run as an Adobe AIR application.



Figure 5: driving simulator

The menu system consisted of 9 top-level items, statically arranged but randomly ordered, each with between 0 and 2 submenus of various sizes. These items consisted of typical in-car functions. A Tobii Glasses Eye Tracker system was employed to gather data on eye movements. This system uses the dark pupil method to measure pupil centered corneal reflection in the right eye at a rate of 30hz, producing information about glance location, timing and duration. To measure subjective workload, the NASA Task Load Index [44] was administered as an electronic form.

4.1.3 Procedure

Participants were introduced to the experiment and signed a consent form followed by being calibrated on the eye tracker. They were then given instructions to change lanes as quickly and efficiently as possible and given approximately 2 minutes to practice, or until they indicated they felt comfortable with the task. Participants were instructed to focus on the primary driving task with 80 percent of their resources, and to use the remaining 20 percent of their resources to execute the secondary IVT task.

This experiment used a 5-level within subject design, in which each participant was exposed to each of four conditions as described in the design section, Direct Touch, Vertical, Circular GRUV and Serial Swipe in a counterbalanced order. One lap of driving on the LCT without an additional secondary task was recorded at the beginning and end of the experiment as a control and between those drives participants had 3 driving blocks per condition, for a total of 14 drives. For each condition, the first block was treated as a training period, and was not analyzed.

Before starting each multi tasking condition, participants were introduced to the device for that condition and asked to practice until they indicated they felt comfortable using it. Participants were not given instructions whether to utilize visual or auditory cues. Once participants passed the start point of the LCT, the experimenter manually initiated the computer's readout of instruction prompts (i.e. find "Navigation" then "Route Preference"). The device would not respond to input until after the command was spoken; after this prompting, the internal timer for task completion would begin. If the participant did not complete the task after 20 seconds, the program would freeze all capability to interact and re-speak the same instructions.

During the drive, deviation from an ideal path was recorded. Once the participant completed all 7 commands for a block, the experimenter ended the drive. This process was repeated for subsequent blocks. After each condition, participants filled out the NASA TLX regarding the previous condition. At the end of all drives, the participant was asked to fill out a demographic and subjective preference survey, given a debriefing, and given a chance to make comments if desired.

4.2 Hypotheses

It was hypothesized that direct touch would lead to significantly larger lane deviation during multitasking compared to the three

multimodal conditions. It was also hypothesized that direct touch would have the fastest average completion time of the secondary task devices, and that the two GRUV conditions would be significantly faster than serial swipe. It was hypothesized that mean dwell time [45] on the secondary task, both in total and as a percentage of multitasking time, would be significantly greater for direct touch compared to the three multimodal conditions. In addition, it was hypothesized that the three multimodal conditions would lead to shorter mean glance times compared to direct touch. Finally, it was hypothesized that cognitive workload would be significantly higher for direct touch than each of the three multimodal conditions.

5. RESULTS

5.1 Primary Driving Task Performance

5.1.1 Planned Analyses

Mean lateral lane deviation while driving was analyzed. Lane deviation data failed the Kolmogorov-Smirnov test of normality ($D=118$, $p<.05$), and exhibited high skewness (.922) and kurtosis (.699). In order to correct for the non-normality, a base 10 log transform was performed to each deviation value, after first adding a constant of 1. A repeated measures ANOVA indicated significant differences amongst the four conditions ($F(3,42)=4.689$, $p<.05$) [Fig.6]. A set of planned simple contrasts compared driving performance for each condition to direct touch ($M=.179$). Only serial swipe led to significantly better driving performance than direct touch ($F(1,14)=5.156$, $p<.05$, $M=.163$), while circle GRUV ($F(1,14)=2.677$, $p>.05$, $M=.201$) and vertical GRUV ($F(1,14)=.006$, $p>.05$, $M=.180$) did not differ significantly from direct touch.

5.2 Secondary Menu Task Performance

5.2.1 Planned Analyses

A repeated-measures ANOVA ($F(3,54)=22.086$, $p<.05$) indicated significant differences in task completion times between conditions. A set of planned simple contrasts indicated that direct touch ($M=4.34s$) was faster than circle GRUV ($F(3,54)=36.143$, $p<.05$, $M=85.64$), vertical GRUV ($F(3,54)=41.311$, $p<.05$, $M=8.14s$), and serial swipe ($F(3,54)=42.062$, $p<.05$, $M=6.85s$). Bonferroni-corrected post-hoc analyses also indicated that serial swipe was significantly faster than circle GRUV ($t(18)=-3.78$, $p<.05$) but not vertical GRUV.

5.2.2 Post-Hoc Analyses

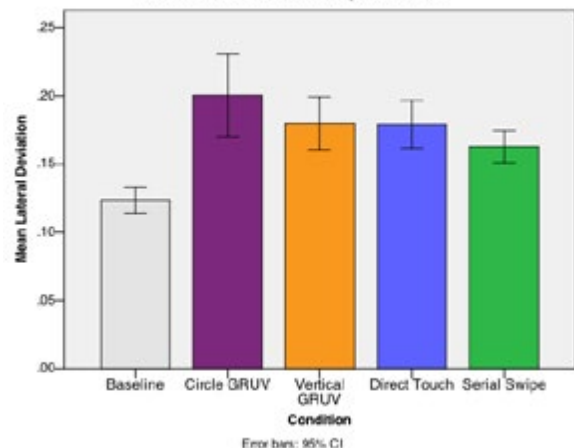


Figure 6: Driving performance by condition.

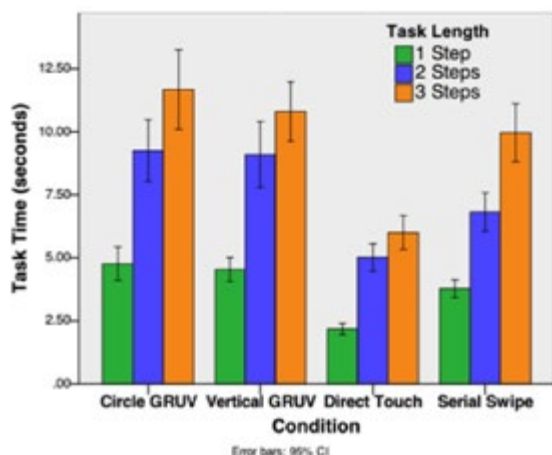


Figure 7: Menu navigation time by number of steps in task.

A Huynh-Feldt repeated measures ANOVA indicated the presence of significant differences in mean task initiation delay for the first task step, between conditions ($F(3,54)=47.787$, $p<.05$). Bonferroni-corrected t-tests indicated significantly higher initiation delay of serial swipe ($M=3.33s$) as compared to circle GRUV ($t(18)=9.69$, $p<.05$, $M=1.84s$), vertical GRUV ($t(18)=10.17$, $p<.05$, $M=1.74s$), and direct touch ($t(18)=8.86$, $p<.05$, $M=2.07s$).

Task time was also analyzed as a function of the number of menu selections needed (between 1 and 3 levels) [Fig.7]. This analysis was unplanned and based on observation during the experiment. Four Huynh-Feldt repeated-measures ANOVAs were conducted.

For circle GRUV, there was a significant effect of target depth ($F(2,36)=30.380$, $p<.05$). Post-hoc analyses indicated significant differences between 1 and 2 level tasks ($t(18)=-7.00$, $p<.05$), as well as 2 and 3 level tasks ($t(18)=2.86$, $p<.05$). For vertical GRUV, the ANOVA indicated differences by target depth ($F(2,36)=40.933$, $p<.05$), with pairwise comparisons showing differences between 1 and 2 level tasks ($t_{18}=-6.44$, $p<.05$), but not 2 and 3 level tasks ($t(18)=2.47$, $p>.05$). This same pattern was observed in the direct touch condition; there was an overall effect of target depth ($F(2,36)=63.607$, $p<.05$), but, while differences between 1 and 2 level tasks were detected ($t(18)=-16.65$, $p<.05$), differences between 2 and 3 level tasks were not ($t(18)=2.42$, $p>.05$). Within the serial swipe condition, there were differences between target depths ($F(2,36)=43.215$, $p<.05$). Post-hoc analyses indicated significant differences between 1 and 2 level tasks ($t(18)=-9.04$, $p<.05$) as well as 2 and 3 level tasks ($t(18)=4.21$, $p<.05$).

A Huynh-Feldt repeated-measures ANOVA indicated no significant differences in error frequency between conditions ($F(3,54)=3.059$, $p>.05$). The observed error rate was highest for circle GRUV (6.3% of selections), and lowest for direct touch (.07% of selections). Error rates for vertical GRUV and serial swipe were 5.5% and 2.8% respectively.

Finally, Bonferroni corrected paired t-tests were conducted comparing block 2 and 3 completion times within each condition. The serial swipe condition was faster in block 3 than in block 2 ($t(18)=2.401$, $p<.05$) by 6.62 seconds. Circle GRUV ($t(18)=1.228$, $p>.05$) and vertical GRUV ($t(18)=1.507$, $p>.05$) were not significantly faster in block 3 than in block 2, potentially due to a large standard error. Participants were faster with direct touch in the third block compared with the second ($t(18)=2.239$, $p<.05$) by 6.51 seconds.

5.3 Glance Behavior

5.3.1 Planned Analyses

A repeated measures ANOVA showed no significant differences between conditions for mean number of task 2 glances ($F(3,39)=1.926$, $p>.05$), mean task 2 glance duration ($F(3,39)=.888$, $p>.05$), or total task 2 dwell time ($F(3,39)=1.566$, $p>.05$).

Glance data was also analyzed as the percentage of dwell time on task 2 during multitasking periods [Fig.8]. A repeated measures ANOVA did not indicate significant differences amongst conditions ($F(3,39)=1.795$, $p>.05$). However, a set of planned simple contrasts comparing each condition to direct touch showed a significant difference in task 2 dwell time percentage between serial swipe and direct touch ($F(1,13)=18.071$, $p<.05$). Direct touch and the two GRUV prototypes were not significantly different in their dwell time percentage.

5.3.2 Post-Hoc Analyses

A glance duration cutoff of 1600ms [34] was applied. Circle GRUV averaged 2.3 such glances, vertical GRUV 4.3, serial swipe 1, and direct touch 3.5. A repeated measures ANOVA showed that the number of glances over 1600ms was not significantly different between conditions ($F(3,15)=.174$, $p>.05$). However, since most participants never glanced for longer than 1600ms, applying this cutoff left very few remaining glances, severely limiting statistical power.

While all glance length distributions were non-normal via KS tests, the direct touch condition exhibited more skewness (2.188 vs. 1.119) and kurtosis (7.461 vs. 1.673) than serial swipe. In addition, the standard deviation of glance lengths was 386.726 for serial swipe, compared with 503.171 for direct touch [Fig.9].

5.4 Subjective Workload

A repeated measures ANOVA was performed on NASA TLX data (including baseline drives). This indicated significant differences between conditions ($F(4,52)=24.024$, $p<.05$). Planned simple contrasts comparing each condition to direct touch indicated that differences between direct touch ($M=57.26$) and baseline ($F(1,13)=29.195$, $p<.05$, $M=26.79$), circle GRUV ($F(1,13)=17.018$, $p<.05$, $M=26.79$) and vertical GRUV ($F(1,13)=7.892$, $p<.05$, $M=65.88$). No significant difference was observed between direct touch and serial swipe ($F(1,13)=1.495$, $p<.05$, $M=50.86$).

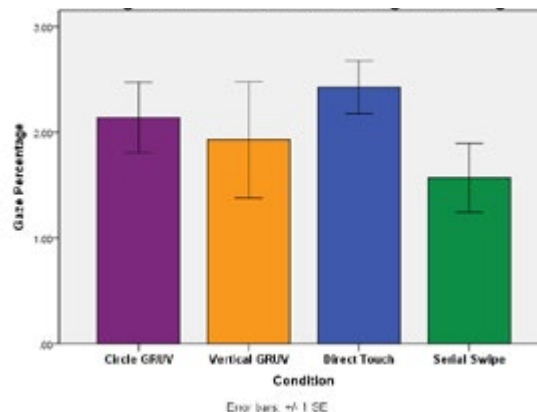


Figure 8: Percent of multitasking time spent looking at device

5.5 Subjective Preferences

The results of the post-experiment questionnaire indicated that participants tended to prefer either direct touch or serial swipe, which accrued 9 and 8 votes respectively. Between the two *GRUV* prototypes, vertical *GRUV* was preferred.

6. DISCUSSION

6.1 Serial Swipe

The serial swipe multimodal condition afforded better driving performance than the direct touch system. While summary eye glance data generally did not support hypotheses about multimodal benefits, the serial swipe condition did have a lower dwell percentage on the menu interface while multitasking, providing a partial explanation for the improved driving performance. The lack of differences in total glance statistics is evidence of the speed and efficiency of direct touch systems—none of the multimodal conditions tended to be used with less overall visual attention. Post-hoc analyses on glance distributions provided no evidence of significant differences, potentially due to the use of the 1600ms cutoff point, which may be high for the simple task of menu selection. Future work could place greater emphasis on planned analysis of glance length distribution features.

There are several possible explanations for why serial swipe conferred driving performance and dwell percentage advantages over direct touch. The first, supported both by the numerical increase in task time across task difficulty for direct touch and through qualitative observation of participant behavior, is that participants resisted taking advantage of the natural breakpoint between steps in multi-level tasks, instead preferring to linger and complete the task in a single attention shift. Further studies could address the magnitude and frequency of this effect. The simplest explanation is that performing tasks with direct touch systems is not effortful or time consuming beyond the initial movement of the hand to the screen, and that the skilled touchscreen user may decide to finish the task with the minimum of time and effort once initially committed. The act of acquiring and tapping a subsequent target when hands (and often eyes) are already on the screen is difficult, and the moment the visual attention shift is extended may be one of dangerous distraction from the roadway. Lee et al. [46] found evidence of a related phenomenon—a tendency for driving performance to degrade and glance length to increase toward the end of long list-browsing tasks using direct touch devices. The present study indicates that similar consideration should be paid to tasks that consist of multiple short sequential steps. Tasks of this sort will become more commonplace as IVTs take on more diverse roles. These findings indicate the necessity of multi-step tasks used for evaluation of vehicle interaction. While driving performance decrement with touchscreens may be minor for one-level selections, the present study indicates that such decrements may increase during multi-step tasks.

Second, the serial swipe condition was much slower than direct touch. Using the serial swipe prototype, due as much to its interaction pattern as latency introduced by multimodal output, was a punctuated, belabored affair, and thus may have afforded frequent return to the primary task [Fig.10]. This is reflected in the fact that the same total dwell time was spread out over a longer period of multitasking. Thus, a degree of built-in slowness might be beneficial in the context of distracted driving. Participants rated serial swipe and direct touch as equally demanding in the NASA TLX, despite the fact that the serial swipe prototype was inherently slower and required a longer sequence of subtasks.

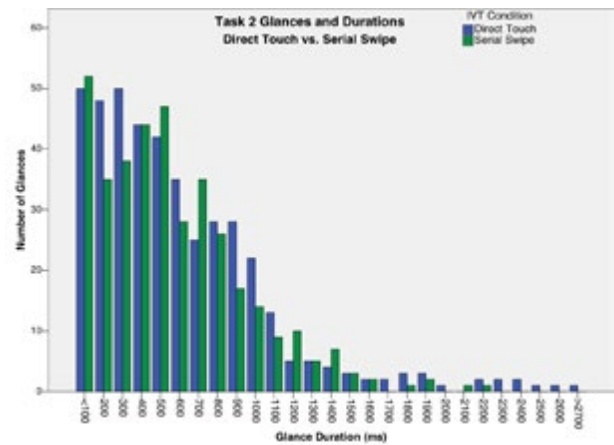


Figure 9: Glance distribution for serial swipe and direct touch

Finally, the design of direct touch does not allow for action that can be easily halted and resumed without loss of progress [Fig.11]. Selecting an item constitutes a long visual search, a goal-directed reaching movement, and potential corrections if a miss occurred. It was observed that errors were sometimes corrected during the same attention shift—further work could evaluate the length of glances immediately following errors in direct touch systems. In contrast, serial swipe allows for incremental progress towards a goal, and for action sequences that can more easily be aborted and resumed. Interruptability and resumability are both important in IVT design [30, 31, 32].

6.2 GRUV

The two *GRUV* conditions were more detrimental to driving performance than both serial swipe and direct touch. In addition, both *GRUV* prototypes, which allow shortcutting and nonlinear exploration, were either slower or the same speed as serial swipe, which only allowed rate-limited sequential movement.

The *GRUV* prototypes were novel, and did not function like anything most participants had experience with. Performance varied widely, and many reported confusion or unmet expectations when using these devices. Participants would periodically become frustrated with *GRUV*—some tried to ‘scroll’ with *GRUV* prototypes (as one would with a traditional trackpad) even after being instructed in how they worked and given practice time. As a whole participants did not appear to memorize menu structures or use shortcuts with *GRUV*, instead preferring to always start at the top and explore downward (vertical *GRUV*) or start at 90 degrees and explore clockwise (circle *GRUV*). Several users did intuit the shortcut pattern, employing occasional quick glancing to plan actions followed by correction using auditory cues. This expert strategy might lead to improved performance with *GRUV* systems.

Alternatively, *GRUV* designs may be unsuitable in the driving context because of their unbroken input/output cycle, regardless of their user’s expertise. Systems like *GRUV* afford a continuous loop [Fig.12] of exploration, evaluation, and further input [39], which reduces the presence of built-in break points to return attention to the road. In addition, if a user aborts an attempt to find a target,

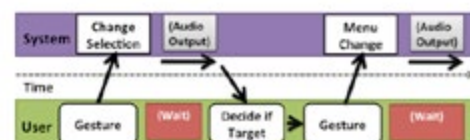


Figure 10: Serial swipe task flow



Figure 11: Direct touch task flow

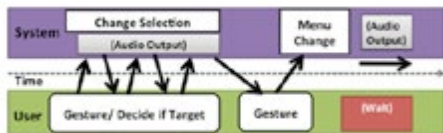


Figure 12: GRUV task flow

they must guess where their finger was on the surface in order to resume where they left off. This often led to either a continuous, unbroken exploration, or a user restarting at the top of the menu.

Finally, *GRUV* conditions may have required substantial fine motor control and spatial awareness- in contrast, using the serial swipe condition only requires finding one large target (the trackpad itself) in which relatively brusque movements can be executed.

6.3 Future Work

It remains to be determined whether it was multimodality or interaction flow that conferred the driving performance advantage for serial swipe over direct touch. This supports the broader thesis that designing multimodal interfaces (especially for IVTs) entails more than adding auditory output to existing systems. Subsequent investigation will attempt to isolate the effects of multimodal output within serial movement devices. Further studies will also attempt to isolate any relationship between the number of steps in a task and the size of performance advantages over direct touch.

While NHTSA guidelines specify that interaction time and step counts should be minimized, they that recommend IVTs should be interruptible without loss of progress and should be operable using many small glances over time [32]. Thus, it might be possible to afford more frequent return to the driving task by rate-limiting selection progress. Doing so without engendering frustration and inefficiency poses a significant design challenge.

Participants were unfamiliar with *GRUV* interfaces. Further work is planned to assess performance given additional exposure to their functionality. It is also important to understand how the present findings may transfer to similar input devices such as turn knobs, wheel-mounted directional pads, or in-air gesture systems. Finally, consideration should be paid to how multimodal, gestural menu systems can be integrated with other IVT interaction methods such as speech input. It may be that a combination of such technologies will most effectively lead to a safer driving experience.

7. REFERENCES

- [1] Newcomb, D. 2010. *The top 10 car technology trends from CES 2010*. Retrieved from <http://www.edmunds.com/ownership/audio/articles/160986/article.html>
- [2] Norman, D. 2002. *The Design of Everyday Things*. Basic Books, New York
- [3] *Auto Alliance Response to NHTSA Voluntary Guidelines to Minimize In-Vehicle Distractions*. 2013. Available from <http://www.autoalliance.org/index.cfm?objectId=526A1BB0-AC46-11E2-9CE9000C296BA163>. Accessed May 2013
- [4] Wickens, C. D., and Seppelt, B. 2002. *Interference With Driving or In-Vehicle Task Information: The Effects of Auditory Versus Visual Delivery*. Technical Report. General Motors Corporation.
- [5] Forlines, C., Schmidt-nielsen, B., Raj, B., Wittenburg, K., and Wolf, P. 2005. A Comparison Between Spoken Queries and Menu-based Interfaces for In-Car Digital Music Selection. in *Proceedings of Human-Computer Interaction - INTERACT 2005: IFIP TC13 International Conference, (Rome, Italy, 2005)*, 536-549.
- [6] Cunningham, Wayne. 2010. *Audi A8 handwriting recognition hands on*. CNet Australia. Available from http://www.cnet.com.au/audi-a8-handwriting-recognition-hands-on_p7-339300398.htm#vp. Accessed February 2013
- [7] Bach, K. M., Jæger, M. G., Skov, M. B., Thomassen, N. G., Vej, S. L., East, A., and Street, B. 2008. You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy April 05-10, 2008) ACM, New York, NY, 1139-1148.
- [8] Wickens, C.D. 1984. Processing resources in attention. In *Varieties of Attention*, R. Parasuraman and D.R. Davies (Eds.). Academic Press, New York, NY. 63-102
- [9] Endsley, M. R. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.
- [10] Chisholm, S.L., Caird, J.K., and Lockhart J. 2008. The effects of practice with MP3 players on driving performance. *Accident Analysis and Prevention*. 40(2), 704-713.
- [11] Neale, V. L., Dingus, T. A., Klauer, S. G., and Goodman, M. 2005. *Overview of the 100-car naturalistic study and findings*. National Highway Traffic Safety Administration, Washington, DC. 1-10.
- [12] Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., and Ramsey, D. J. 2006. *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-Car Naturalistic Driving Study Data*. National Highway Traffic Safety Administration, Washington, DC.
- [13] Blanco, M., Biever, W. J., Gallagher, J. P., and Dingus, T. 2006. The impact of secondary task cognitive processing demand on driving performance. *Accident analysis and prevention*, 38(5), 895-906.
- [14] Pitts, M. J., Skrypchuk, L., Wellings, T., Attridge, A., and Williams, M. 2012. Evaluating User Response to In-Car Haptic Feedback Touchscreens Using the Lane Change Test. *Advances in Human-Computer Interaction*, 2012, 1-13.
- [15] Ranney, T.A, Harbluk, J.L., and Noy, Y.I. 2002. The effects of voice technology on test driving performance: Implications for driver distraction. In *Proceedings of the Human Factors and Ergonomics 46th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society. 1814-1818.
- [16] Lee, J., and Spence, C. 2008. Assessing the Benefits of Multimodal Feedback on Dual-Task Performance under Demanding Conditions. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*. (Liverpool, United Kingdom, 1-5 September 2008). British Computer Society, Swinton, UK, UK. 185-192.
- [17] Srinivasan, R., Yang, C., Jovanis, P., Kitamura, R., Anwar, M. 1994. Simulation study of driving performance with selected route guidance systems. *Transportation Research*, 2C(2), 73-90.
- [18] Hurwitz, J., and Wheatley, D.J. 2002. Using driver performance measures to estimate workload. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*. 1804-1808.

- [19] Walker, J., Alicandri, E., Sedney, C., and Roberts, K. 1991. In-vehicle navigation devices: Effects on the safety of driver performance. In *Vehicle Navigation and Information Systems Conference Proceedings*. (October 20-23 1991) Society of Automotive Engineers, Warrendale, PA 499-525
- [20] Dingus, T.A., Hulse, M.C., McGehee, D.V., and Manakkal, R. 1994. Driver performance results from the Travtek IVHS camera car evaluation study. In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society*. Human Factors Society. Santa Monica, California. 1118-1122
- [21] Sodnik, J., Dicke, C., Tomažič, S., and Billingham, M. 2008. A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies*, 66(5), 318–332.
- [22] Moskovitch, Y., Jeon, M., and Walker, B. N. 2010. Enhanced Auditory Menu Cues on a Mobile Phone Improve Time-Shared Performance of a Driving-Like Dual Task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (San Francisco, USA. September 27-October 1, 2010). 54(18), 1321–1325.
- [23] Janssen, C. P., and Brumby, D. P. 2010. Strategic adaptation to performance objectives in a dual-task setting. *Cognitive Science*, 34(8), 1548–60.
- [24] Jeon, M., Davison, B. K., Nees, M. A., Wilson, J., Walker, B. N., and Street, C. 2009. Enhanced Auditory Menu Cues Improve Dual Task Performance and are Preferred with In-vehicle Technologies. *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. (Essen, Germany. September 21-22 2009). ACM, New York, NY. 91–98.
- [25] Liu, Y. C. 2001. Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, 44(4), 425–42.
- [26] Parkes, A.M., and Burnett, G.E. 1993. An Evaluation of Medium Range "Advance Information" in Route-Guidance Displays for Use in Vehicles. *IEEE Vehicle Navigation and Information Systems Conference*, Ottawa, Canada. 238-241
- [27] Helleberg, J. and Wickens, C.D. 2002. The effects of data link modality and display redundancy on pilot performance: an attentional perspective. *The International Journal of Aviation Psychology*. 13(3) 189-210
- [28] Recarte, M.A. and Nunes, L.M. 2000. Effects of verbal and spatial- imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*. 6 31–43.
- [29] IHS Electronics and Media. 2012. *Automotive Touchscreen Market Booming*. Press Release.
- [30] Green, P. 2000. Crashes Induced by Driver Information Systems and What Can Be Done to Reduce Them. In *Automotive electronics : delivering technology's promise : proceedings of the 2000 International Congress on Transportation Electronics*. (Detroit, Michigan, USA. October 16-18, 2000). Society of Automotive Engineers, Warrendale PA. 27-36
- [31] Driver Focus-Telematics Working Group. 2012. Statement of Principles, Criteria and Verification Procedures on Driver-Interactions with Advanced In-Vehicle Information and Communication Systems. Alliance of Automobile Manufacturers, Washington, DC.
- [32] U.S. Department of Transportation. 2012. *National Highway Traffic Safety Administration*. Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices.
- [33] Fitts, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391
- [34] Horrey, W.J., and Wickens, C.D. 2007. In-vehicle glance duration: distributions, tails, and model of crash risk. *Transportation Research Record*. 2018, 22-28.
- [35] Rydström, A., Broström, R., and Bengtsson, P. (2012). A comparison of two contemporary types of in-car multifunctional interfaces. *Applied ergonomics*, 43(3), 507–514.
- [36] Ecker, R., Broy, V., Hertzschuch, K., and Butz, A. 2010. Visual Cues supporting Direct Touch Gesture Interaction with In-Vehicle Information Systems. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. (Pittsburgh, Pennsylvania, USA. November 11-12 2010). ACM, New York, NY. 80-87.
- [37] Weinberg, G., Knowles, A., and Langer, P. 2012. Bullseye : An Automotive Touch Interface that 's Always on Target. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. (Portsmouth, New Hampshire, USA, October 17–19.) ACM, New York, NY.
- [38] Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., and Baudisch, P. 2007. earPod : Eyes-free Menu Selection using Touch Input and Reactive Audio Feedback. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA April 38- May 03, 2007) ACM, New York, NY. 1395- 1404
- [39] Zhao, S., Brumby, D. P., Chignell, M., and Salvucci, D. 2013. Shared Input Multimodal Mobile Interfaces : Interaction Modality Effects on Menu Selection in Single-Task and Dual-Task Environments. *Interacting With Computers*
- [40] Kane, S. K., Bigham, J. P., and Wobbrock, J. O. 2008. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-Touch Interaction Techniques. *Proceedings of the Tenth International ACM SIGACCESS Conference on Computers and Accessibility* (Halifax, Nova Scotia, Canada, October 13 - 15, 2008). 73–80.
- [41] Yalla, P., and Walker, B. N. 2008. Advanced auditory menus: Design and evaluation of auditory scroll bars. *Proceedings of the Tenth International ACM SIGACCESS Conference on Computers and Accessibility*, (Halifax, Nova Scotia, Canada, October 13 - 15, 2008). ACM, New York, NY. 105-112.
- [42] AT&T Research Lab, *AT&T Text-to-Speech Demo*. Available from <http://www.research.att.com/projects/tts/demo.html>.
- [43] Mattes, S. 2003. The lane change task as a tool for driver distraction evaluation. In *Quality of work and products in enterprises of the future*, H. Strasser, H. Rausch, and H. Bubb (Eds). Ergonomia, Stuttgart, Germany. 57
- [44] Hart, S. G., and Staveland, L. E. 1988. Development of NASATLX (Task Load Index): Results of empirical and theoretical research. In A. H. Peter and M. Najmedin (Eds.), *Advances in psychology* 52(1) North-Holland, Amsterdam, The Netherlands. 139–183
- [45] Industry Standards and Regulations. 2000. SAE J2396: Definitions and Experimental Measures Related to the Specification of Driver Visual Behavior using Video Based Techniques.
- [46] Lee, J. D., Roberts, S. C., Hoffman, J. D., and Angell, L. S. 2012. Scrolling and Driving: How an MP3 Player and Its Aftermarket Controller Affect Driving Performance and Visual Behavior. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(2), 250–263.