INVESTIGATING HUDs FOR THE PRESENTATION OF CHOICE LISTS IN CAR NAVIGATION SYSTEMS

Garrett Weinberg[1], Bret Harsham[2], Zeljko Medenica[3]
[1,2]Mitsubishi Electric Research Labs, [3]University of New Hampshire
[1,2]Cambridge, MA, USA, [3]Durham, NH, USA
[1]garrett.weinberg@nuance.com (now with Nuance Communications), [2]harsham@merl.com, [3]zeljko.medenica@unh.edu

Summary: It has been established that head-down displays (HDDs), such as those commonly placed in the dashboard of commercial automobiles, can draw drivers' attention away from the primary driving task (Bach et al., 2008). This problem can be exacerbated when screens are "busy" with graphics or rich information. In this paper, we present the results of a driving simulator-based user study where we examined two potential alternatives to HDDs for presenting textual lists. Subjects conducted a series of street name finding tasks using each of three system variants: one with a HDD, one with a head-up display (HUD), and one with only an auditory display. We found that the auditory display had the least impact on mental load, but at the expense of task completion efficiency. The HUD variant also had a low impact on mental load and scored highest in user satisfaction, and therefore appears to be the most viable target for future study.

INTRODUCTION AND MOTIVATION

Head-up displays are partially-transparent displays that render information in a manner that allows the viewer to comprehend it while remaining focused on the scene in front of her. General Motors introduced the first automotive HUD in 1988 (Weihrauch et al., 1989), and HUDs have since been deployed on a variety of vehicles from a variety of manufacturers, though mostly in the luxury segment and mostly as costly optional upgrades. They typically display only limited, critical information such as vehicle speed and/or the direction of an upcoming turn (if route guidance is active). The aim of this paper is to examine whether HUDs might also be appropriate for non-critical textual information, such as the destination- or media-related menus or lists that today are presented on head-down displays (HDDs).

With the proliferation of both built-in and nomadic devices used in the car and the increasing amount of rich media content available on them, there is a strong practical motivation to find safer ways of presenting this content. As we have seen with mobile phone use, just because an activity has been proven unsafe (e.g., Horrey et al., 2006), that doesn't mean drivers will stop engaging in that activity. Thus, it is imperative that we develop displays and user interfaces that allow access to the ever-widening array of "infotainment" with as little adverse effect on driving as possible. As a foray into this research area we designed a driving simulator-based user study that compared subjects' task performance, preferences, and reported workload when carrying out a typical in-vehicle information system (IVIS) task—street name retrieval—using three variants of a prototype IVIS: one with a HUD, one with a HDD, and one with no display at all.

RELATED WORK

Numerous studies have illustrated the benefits of HUDs for the presentation of information related to the operation of the vehicle itself. For example, Sojourmer & Antin (1990) found that subjects operating a driving simulator that featured an HUD speedometer reacted significantly more quickly to salient cues in the driving scene versus a typical dashboard-mounted speedometer. Burnett (2003) found that participants made fewer navigational errors (wrong turns) when an arrow graphic and contextual information about the surrounding roads was presented on a HUD versus on a HDD. Charissis et al. (2008) used a simulator to test an HUD-based collision warning system in low-visibility conditions (typical of heavy rain or fog). They found that the HUD dramatically reduced the number of collisions and improved subjects' maintenance of following distance, when compared to a traditional HDD.

Less research has been done on drivers' ability to process information *not* immediately related to vehicle operation, such as music titles or points of interest, when such information is presented on HUDs. Often this kind of information is best conveyed by textual descriptions, or some combination of textual, visual, and audible descriptions. Tsimhoni et al. (2001) investigated the best position for presenting short text messages on a full-windshield HUD, evaluated in terms of driving performance and workload. They found that message locations within 5° of a straight-ahead gaze yielded the best performance and were preferred by subjects. Note, however, that these conclusions were obtained based on very short messages only (people's names), as was also the case with (Ablassmeier et al., 2007) (where the messages were simple navigation instructions). Our work seeks to determine whether such findings will hold for a multi-line, information-rich HUD.

EXPERIMENT

We conducted an experiment in which subjects were asked to carry out choice list navigation tasks (finding specific street names) using both a HUD-based IVIS and two "baseline" variants: one with a HDD and one without any visual display at all—only synthesized speech output.

Driving Simulator: The experiment was conducted in a mid-range driving simulator that incorporates motion feedback in the form of 2-axis tilt and vibration that varies with vehicle speed and road texture. For more information on the simulator's software and hardware configuration, please see (Weinberg & Harsham, 2009). The virtual roadway used in this experiment was the same one used in (Weinberg et al., 2010): a gently curving highway course with a single leading vehicle that travels at approximately 90 km/h (55 mph) on straight portions but slows to as little as 48 km/h (30 mph) on curved portions.

Prototype IVIS

The prototype voice-activated IVIS with which subjects interacted during the study was similar to that used in (Weinberg et al., 2010). There was a main menu representing three domains of content: Navigation, Music, and Contacts, and each of these domains contained various subdomains and functions (some of which were mocked up). Users could traverse the hierarchy

either by issuing voice commands or by selecting and clicking using a vertically-oriented "jog dial" device mounted on the steering wheel.

Choice List Behavior. Whereas other systems will announce the top several choices one after another and then immediately open the microphone for input (e.g., the original Ford Sync), our system maintains a logical cursor that the user controls with the steering wheel-mounted jog dial device. Each entry is announced—prefaced by its number—only when the cursor reaches that entry. To actually activate an entry, the user must either press the jog dial's central momentary switch while the cursor is on that item, or press the push-to-talk button and say the number of any item.

HDD Variant. The head-down variant of the IVIS was positioned near the steering wheel as illustrated on the left in Figure 1, at a slight upward tilt to allow for easy reading. With the simulator's seat in its rearmost position (as it was for the tallest of subjects), the center of the display was 30° right of center, approximately 23° below the line of sight for a 1.75m (5 ft. 9 in.) person. It occupied about 11 horizontal and about 8 vertical degrees of the field of view.
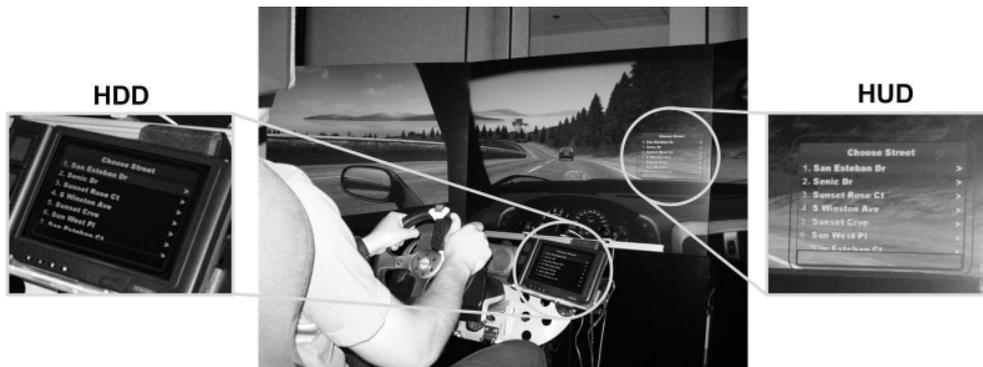

Figure 1: HDD (left) and HUD (right) variants of our prototype IVIS (shown together for illustrative purposes; only one at a time was used in the experiment)

HUD Variant. The head-up variant was superimposed on the driving scene using a video overlay mixer. The original black background of the IVIS interface was removed by the mixer, thus creating an impression of the interface being semi-transparent. Only the choice list text and the selection bar remained fully opaque. As shown on the right of Figure 1, the HUD was positioned 14° to the right of center and approximately even with the vertical line of sight. It occupied about 8 horizontal and 7 vertical degrees of the field of view. This positioning allowed subjects to glance away from the driving scene as often as necessary in short bursts, without changing their depth of focus. The effective focal distance in our simulator is shorter than that of real, three-dimensional driving, but we believe this is a fair representation of real-world HUDs, because they are designed such that their image appears to the viewer to be within the driving scene (at bumper depth or beyond).

Audio-Only Variant. The audio-only variant behaved identically to the HDD and HUD variants, but simply lacked a visual display completely. This forced subjects to rely solely on the status messages that the system generated via its text-to-speech (TTS) engine. To be clear, these message outputs were present in all three variants; however in the audio display variant they were the *only* form of output.

Experimental Protocol

24 subjects, 10 female and 14 male (age M=22.45, SD=3.64), were recruited from the local area to participate in our study. Subjects were required to be licensed drivers. Each experimental session was about one hour and thirty minutes long, and subjects were compensated $40 for their participation. The experiment consisted of four drives: control, head-up, head-down, and audio-only. In the control drive the subjects operated the simulator without any IVIS interactions, whereas in the other drives, they interacted multimodally (voice + manual controls) with the given IVIS variant while simultaneously driving. Each drive lasted for seven minutes, and all subjects completed all four drives. All drives took place on the same simulated roadway. To account for learning effect, we counterbalanced the order of the drives among subjects. The main task in all four drives was to follow a pace vehicle and to maintain a constant distance behind it. Subjects were instructed that following the vehicle and driving safely had the highest priority, while all other tasks (i.e., listening to the experimenter's prompts and operating the IVIS) had secondary importance.

Before starting the experiment itself, subjects had a five minute training period to get accustomed to the driving simulator. Before the first street-retrieval (i.e., non-control) condition, the subjects were trained on how to operate the IVIS until they were comfortable searching for street names on their own (the baseline HDD variant was always used for this training).

Street Finding Tasks. During the three IVIS drives, the experimenter prompted subjects to find various randomly-chosen streets from a database of all non-numerically named streets in California. The subjects began each interaction by saying "Enter destination" followed by the target street name. Thereafter, subjects were instructed to use whichever combination of voice commands and manual controls they found easiest; however, there was no purely-manual affordance for entering the street names themselves, only voice. The same voice dialog design was used in all three conditions—only the means of presenting the result choice list varied.

Each of the first 7 attempted street retrievals was considered to be a separate trial for the purposes of our repeated-measures analysis (including multiple subsequent attempts to find the same street). The time between attempts to find different streets was not less than ten seconds. This allowed the subject some down-time to resume following the lead vehicle at the preferred distance. A task was considered to be successfully completed when the sought item was found and activated.

The position of the target item within the list was artificially controlled; in other words, the actual output of the speech recognizer was ignored in the street name phase of the voice dialog. 50% of the time, the target street name appeared at the top of the list, 25% of the time in a random position elsewhere on the first page (entries 2 – 7), and 20% of the time in a random position on the second page of results (entries 8 – 14). 5% of the time, the target item did not appear on the list at all, forcing the subjects to issue the "*Enter destination*" command again (at which point the target item was re-shuffled into the list according to the given distribution). The non-target entries on the list were filled with similar sounding street names chosen automatically using the Soundex algorithm (Knuth, 1968), in order to maintain the appearance of real ASR (albeit unrealistically inaccurate ASR) while controlling for error rate.

Following the completion of each drive, subjects were instructed to complete the NASA-TLX survey (Hart & Staveland, 1988) considering the driving task alone or the combination of driving and retrieval tasks, as appropriate. At the end of the entire experiment, a user interface preference questionnaire was administered. In addition to task timing and workload (NASA-TLX) data, we also captured eye glance durations and a variety of driving behavior metrics. We omit these results here for space reasons but will report them in a later work.

IVIS Usability Results

Task Time. Task time was defined as the amount of time which elapsed between the beginning and the end of a task. The beginning of each task was marked at the end of the experimenter's task prompt (e.g., "Please find Sunset Boulevard"). As for the end of the task there were a few possibilities: the participant selected the desired street name, declared the current task to be a failure, or issued a new "Enter destination" command to start a new search.
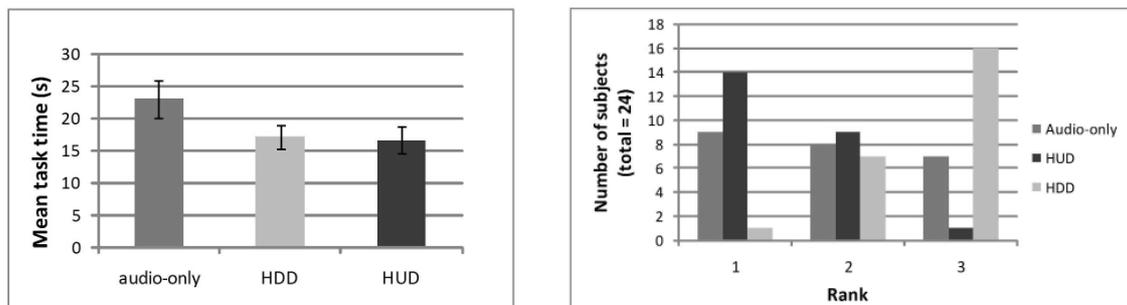

Figure 2. Left: Mean task times for the three *IVIS-variants* (error bars represent 95% confidence intervals) Right: Cumulative rankings of the three *IVIS-variants (1 is highest)*

There was a significant main effect of the *IVIS-variant* on *task time* ($F_{2,46}$ = 17.250, p < 0.001). The mean task times for audio-only, HDD, and HUD conditions are 22.9s, 17.1s, and 16.6s, respectively (Figure 2L). A post-hoc comparison indicated significant differences between audio-only and HDD (p < 0.001) and audio-only and HUD (p < 0.001). No difference was observed between HUD and HDD (p = 0.547).

Table 1. Subjective opinions about *IVIS-variants*. Number of subjects who fully agreed (FA), agreed (A), fully disagreed (FD), or disagreed (D) with each statement

Statement	Agreement	Audio-only	HDD	HUD	p (χ^2)
The ___ interface was easy to use.	FA/A	19	15	23	**0.003** (11.4)
	FD/D	4	4	1	
The ___ interface distracted me from driving.	FA/A	3	19	8	**0.001** (27.9)
	FD/D	17	2	7	

Subjective Opinions and Rank. After concluding the study, each participant had a chance to express her subjective opinion about each of the *IVIS-variants* using a Likert scale questionnaire. Table 1 shows the number of participants and their agreement with statements about ease-of-use and driving distraction for each interface. The statistics for the main effect of the *IVIS-variant* are obtained using the Friedman non-parametric test. The rightmost column in Table 1 illustrates a significant main effect of *IVIS-variant* for both statements. Post-hoc comparisons were

conducted for both questions using the Wilcoxon Signed Rank test and Bonferonni correction for multiple comparisons. The results revealed the following significant differences: for ease-of use – HDD vs. HUD (p = 0.003) and for distraction from driving – audio-only vs. HUD (p = 0.006), audio-only vs. HDD (p < 0.001), HDD vs. HUD (p < 0.001). At right in Figure 2 we see how subjects ranked the interfaces with respect to the desirability of using them in their own vehicles (1: most desirable, 3: least desirable). The results show a significant main effect of the IVIS-variant on the subjective ranking (χ^2=16.583, p < 0.001). A post-hoc comparison revealed significant differences for audio-only vs. HDD (p = 0.033) and HDD vs. HUD (p < 0.001).

Subjective Mental Load Evaluation. Immediately after completing each of the driving sessions, subjects filled out a NASA-TLX questionnaire in order to assess their workload while interacting with each of the interfaces. Figure 3 shows the mean reported workload for each condition. We found a significant main effect of the IVIS-variant on mental load ($F_{3,69}$ = 16.484, p < 0.001). The post-hoc comparison showed highly significant differences (p < 0.008) for all parings except audio vs. HUD (p = 0.699).
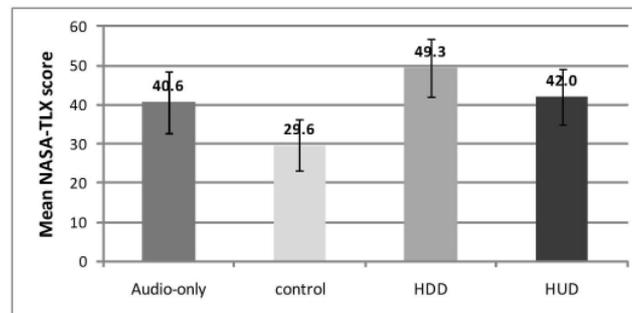

Figure 3. Mean workload values for each IVIS-variant obtained through NASA-TLX (higher values indicate higher mental load)

Discussion

The results show that HUD was the most preferred by participants, although not significantly so in comparison to audio-only. Even though the audio-only variant imposed a marginally lower cognitive load than HUD, the fact that its tasks took so much longer might explain why it did not achieve the highest subjective rank. Participants' relative dislike of the HDD variant (Figure 2R) may stem from its higher reported workload (Figure 3), since its task completion time is on par with that of the HUD (Figure 2L). Participants' preference for the HUD variant (Table 1) may be founded upon this interface's perceived ease-of-use and its relatively low required task time (Figure 2), because it did not win out in terms of subjects' opinions about the system's distraction potential. This metric shows the audio-only variant to be the most favorable. Subjects may have preferred the HUD in spite of its greater impact on driving because of its novelty, or because its visual modality afforded a working memory "crutch" that was missing in the purely auditory variant.

CONCLUSIONS AND FUTURE WORK

Based on the presented results we can conclude that displaying choice lists using auditory displays or head-up displays are promising techniques. If we consider workload alone, audio-

only is the preferred approach. However, if we broaden our focus to consider user satisfaction and task efficiency as well, HUD appears to be the best alternative. Future work in this area should seek a middle ground that offers the attention and safety benefits of an audio-only approach but allows tasks to be completed as efficiently as they can with a multimodal HUD.

One potential way forward would be to augment the audio-only interface with a "sometimes-on" HUD. The HUD would only be displayed when a content-retrieval task is active, and would slide out of view after a short interaction timeout. Alternatively, one could display the HUD choice list only in situations when the recognition accuracy is poor (indicated by a low score from the ASR engine). In any case, further validation and refinement of these display techniques will be necessary if they are to displace—or augment—the ubiquitous head-down LCD screens in real production vehicles.

REFERENCES

Ablassmeier, M., Poitschke, T., Wallhoff, F., Bengler, K., and Rigoll, G. (2007). Eye Gaze Studies Comparing Head-Up and Head-Down Displays in Vehicles. *Proceedings of the IEEE Int. Conf. on Multimedia and Expo*, Beijing, China, 2250-2252.

Bach, K., Jæger, M., Skov, M. B., and Thomassen, N. (2008). You can touch, but you can't look: interacting with in-vehicle systems. *Proceedings of the International Conference on Human Factors in Computing Systems (CHI)*. Florence, Italy, 1139-1148.

Burnett, G. E. (2003). A road-based evaluation of a head-up display for presenting navigation information. *Proceedings of the HCI International Conference*. Crete, Greece, 3, 180-184.

Charissis, V., Papanastasiou, S., and Vlachos, G. (2008). Comparative study of prototype automotive HUD vs. HDD: collision avoidance simulation and results. *SAE World Congress & Exhibition*. Detroit, MI.

Hart, S. G. and Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. *Human Mental Workload.* 139-183.

Horrey, W.J. and Wickens, C.D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48* (1), 196-205.

Knuth, Donald Ervin. (1968). *The art of computer programming.* Reading, MA: Addison-Wiley.

Sojourmer, R. J. and Antin, J. F. (1990). The effects of a simulated head-up display speedometer on perceptual task performance. *Human Factors, 32* (3), 329-339.

Tsimhoni, O., Green, P., and Watanabe, H. (2001). Detecting and reading text on HUDs: Effects of driving workload and message location. *Proceedings of the IEEE 4th International Conference on Intelligent Transportation Systems.* Oakland, CA.

Weihrauch, M., Meloeny, G.G., and Goesch, T.C. (1989). The first head-up display introduced by General Motors. *SAE Technical Paper 890288.*

Weinberg, G. and Harsham, B. (2009). Developing a low-cost driving simulator for the evaluation of in-vehicle technologies. *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI).* Essen, Germany, 51-54.

Weinberg, G., Medenica, Z., Harsham, B., and Forlines, C. (2010). Contextual push-to-talk: shortening voice dialogs to improve driving performance. *Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI).* Lisbon, Portugal, 113-122.