IDENTIFICATION OF AN "APPROPRIATE" DROWSY DRIVER DETECTION INTERFACE FOR COMMERCIAL VEHICLE OPERATIONS

Ellen M. Ayoob
Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pa

Aaron Steinfeld
Robotics Institute
Carnegie Mellon University
Pittsburgh, Pa

Richard Grace
Attention Technologies
Pittsburgh, Pa

Considerable progress has been made in measuring drowsiness and understanding its effects upon human performance in the laboratory and in simulated and operational driving conditions. This work builds upon previous research and identifies an appropriate design for a drowsy driver detection interface. A participatory design process was used that included both design experts and drivers in separate focus groups. One expert activity, evaluations of candidate interaction flow models, and two driver activities, critical incident interviews and a design exercise, are described here. The conflict that arose between the drivers' desires and the desires of the scientific community is that the drivers viewed the system as a loyal servant that would alert the driver when he became drowsy, while the scientific community viewed the system as a trusted advisor that would encourage the driver to stop and rest. The final design has many features to address both of these views.

INTRODUCTION

Driver drowsiness poses a major threat to roadway safety and the problem is particularly severe for commercial motor vehicle (CMV) drivers. Drowsy driver crashes cost $12 billion and contribute to up to 35% of the 4,400 annual truck driver deaths (FHWA, 1998). Fatigued drivers are often unaware of their condition, frequently driving for 3-30 seconds with their eyes closed. Twenty-four hour operations, high annual mileage, exposure to demanding environmental conditions and demanding work schedules make drowsiness a major cause of combination-unit truck (CUT) crashes.

Considerable progress has been made in measuring drowsiness and understanding its effects upon human performance in the laboratory and in simulated and operational driving conditions. Wierwille, et al (1994) generated a measure of drowsiness, PERCLOS, associated with degradation in driving performance in a simulated roadway environment. Experimental studies performed by Dinges, et al (1998) to test the validity of PERCLOS and other new technologies for drowsiness detection showed that PERCLOS was able to accurately predict fatigue-induced lapses in vigilance. Studies by Grace, et al (1999) of overnight commercial trucking operations have produced a real-time monitor capable of detecting driver drowsiness in an operational setting. Furthermore, this monitor used in conjunction with a driver feedback system has been shown to decrease drowsiness and improve driver performance in simulated driving conditions (Mallis et al, 2000). These advances, for the first time, make accurate detection and management of drowsiness feasible.

The main question to the driver is which of the driving time countermeasures and rest time countermeasures are desirable, practical and/or useful. It is known that countermeasures other than sleep may be ineffective or effective for only short periods of time (Mallis et al, 2000). Despite this, anecdotes or myths about personal habits may instill drivers with false confidence about the effectiveness of their personal method. It is also common knowledge within the scientific community (Dinges, 1989; Wylie, 1996; Brown, 1997) that self-assessment of drowsiness is unreliable. A driver may decide to disregard the warning from the feedback system based on his/her own perception.

User-centered interfaces and corresponding interactions for warning systems should enable drivers to understand the severity of the warning and adjust their behavior accordingly. As such, a major research question for this effort was:

- What is the most appropriate design for a drowsy driver detection interface?

In particular, drivers need accurate, reliable information that highlights their condition without providing them with encouragement to extend unsafe behavior or contributing to the perception that repeated warnings are a substitute for sound judgment. Modification of behavior with safer choices should reduce the frequency of highly fatigued periods, and lead to a decrease in potential incidents.

DESIGN APPROACH

The design process was organized around humanistic themes (human connection, choice, engagement, integration, driver awareness, and association) that emerged from gaps identified in past approaches. These themes set the tone for the designs.

With user-centered design principles (Jordan, 1998) central to the development, the team undertook a series of activities to enhance their understanding of the problem space and assist in the development of an appropriate design. These included literature reviews, brainstorming sessions, field visits, and thematic explorations that culminated in an expert/advisor focus group and a user focus group.

By talking with design and usability experts first, less usable or conceptually flawed ideas were eliminated and valuable knowledge was gained. By speaking with a representative group of drivers, their perceptions, preferences, issues, and attitudes became known. Unfortunately, constraints did not permit more than one user focus group.

Focus Group 1 – Design Experts

This session contained elements of a structured, cognitive walkthrough. Design and human factors experts reviewed potential interaction concepts and discussed the meaning and consequences of "appropriateness" as the defining design factor. There were six activities total: short context-setting film, tour of a truck cab, high-level briefing of goals, interaction flow model exercise and design review, a random spot signaling exploration, and a sound exploration. The evaluations of candidate interaction flow models were particularly interesting with respect to the findings and will be focused on here.

Four interaction flow models were presented to the experts for their critical review. The interaction flow models focused on the human-to-machine interaction rather than machine-to-machine interaction. Every user action or inaction that results in a system state change, and how they would work through the system, was presented from the users' perspective.

The interaction flow models were presented as large printouts accompanied by blue foam models (Figure 1), printouts of simulated interfaces, and a list of discussion questions. The foam models and simulated interfaces were all vastly different from concept to concept, intentionally, so that the experts would understand that the aesthetics were secondary and variable and focus on the human-device interaction instead. These models were:

- *Model One, Audio Only,* was adapted from an existing prototype developed by Grace and Stewart (2001) where the primary interaction for the user was a tone and a bank of LEDs. The interaction of the prototype was varied to add the option of user-selected sounds rather than just one tone.
- *Model Two, Hierarchy,* was adapted from a sequence proposed by Wierwille, Lewin, & Fairbanks (1995) that featured countermeasures that are accessible through a hierarchy of use. A user would not experience underlying functionality unless they progressed to that point in their drowsiness. The model was adapted to introduce additional countermeasures that focus on human connection at the point of severe drowsiness, such as an alarm triggering a phone call to a designated support person who could speak with the driver.
- *Model Three, Sensitivity,* was a new concept that introduced the idea of a user controlling the sensitivity of the system based on their drowsiness state. This sensitivity setting would be connected to the rate of alarms and was thought to be potentially helpful with user annoyance and tolerance of false alarms. The main point, though, was that the driver would set the unit based on how alert they were feeling.
- *Model Four, Modality,* was a new concept that introduced the idea of a user being able to specify a modality preference, such as a visual, auditory and/or possibly vibratory warning, depending on the user's needs or preferences from drive to drive.

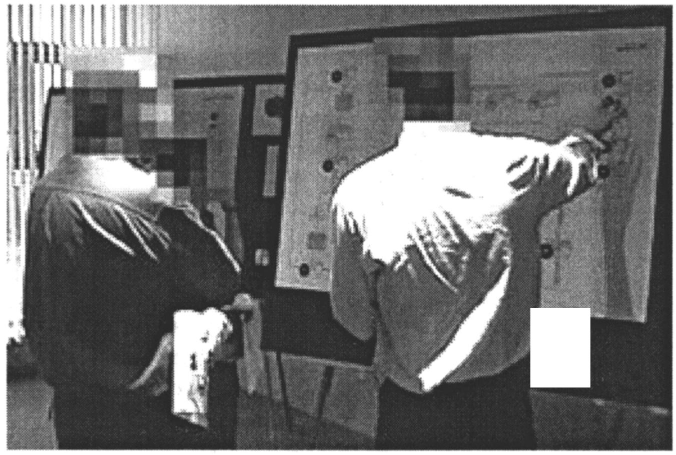

Figure 1: Experts reviewing an interaction flow.

Focus Group 2 – CMV Drivers

The main goal of this focus group was to collect information from less-than-load (LTL) drivers about their perceptions, preferences, resistance, attitudes, acceptability, and the consequences surrounding the adaptation and use of a drowsy driver detection and warning system. There were four activities total: questionnaire, action sequence model (Beyer, 1998), critical drowsy driving incident interviews, and a design exercise and subsequent group discussion. The critical incident interviews and design exercise will be focused on here. All of the activities were structured so that the participants would become increasingly invested in the interaction decisions they were going to be asked to make during the design exercise.

The critical incident technique developed by Flanagan (1954) is a way to collect task-based data from interviews by asking the interviewees first to recall a specific critical incident (story of a real situation) then answer a set of prepared questions to reveal more information about the specific incident (as explained in Hackos, 1998).

An experimenter interviewed each driver alone in order to reduce confidentiality concerns and encourage honest answers. The purpose was to collect anecdotal evidence about each driver's experience with having a drowsy driving incident, whether it was a near miss or an actual crash, in order to detect patterns of behavior. These patterns of behavior, while statistically lacking, are common enough to provide clues to the resistance and/or acceptance of a drowsy driver detection and warning system. It was also thought that this exercise, as a precursor to the design exercise, would help set the context and tone and reinforce the seriousness of the issue.

The purpose of the design exercise was to explore user preferences, the context of use, and usability issues within teams of two first and then as a group last. By dividing the drivers into teams of two, they were encouraged to communicate and reach consensus on their choices. It was also a way to negate any interpersonal dynamics that may have developed in the focus group during the other activities that were mentioned earlier.

It was decided to present the design choices as descriptions rather than low fidelity visual prototypes to keep the discussion and choices as broad based as possible. The choices were extracted from the interaction flow models that were presented to the experts and refined based on their comments. For example, the drivers could choose whether they would turn the system on manually or have it power on automatically via the ignition switch. A more subjective example is whether they would want the display to operate like a gas gauge, a check oil light, etc., and what types of information they might want to see. Each team was given a kit that contained preset choices and a work board where they could affix them and annotate comments (Figure 2). They also had the opportunity to hear sounds and experience a vibrating seat to help inform their decision-making process.

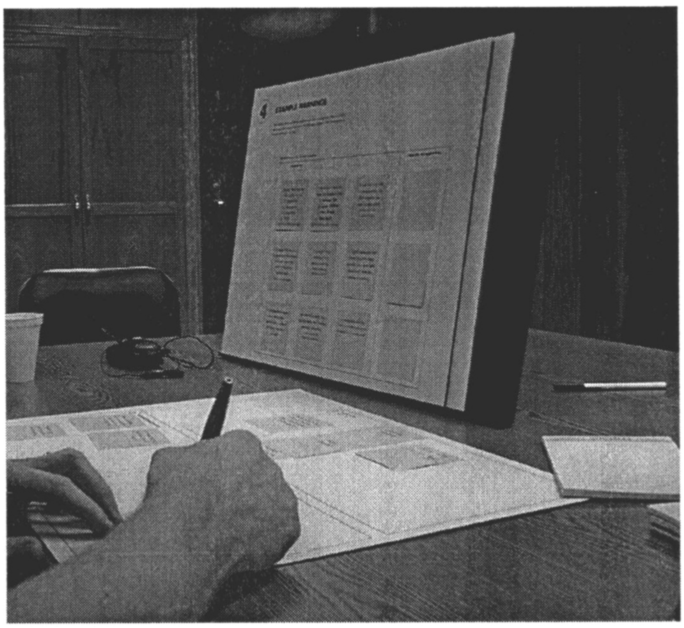

Figure 2: Driver participating in the design exercise.

RESULTS

Focus Group 1 – Design Experts

The major comments and concerns with the interaction flow models addressed issues such as:

- Whether the device should be informational, diagnostic, monitoring, or a snooze alarm
- The manner in which the device turns on
- The system being smart enough to learn or adapt over time
- Showing drowsiness information in an explicit manner whether drivers want to collect information about themselves

Specific comments about the interaction flow models are paraphrased in Table 1.

Table1. Summarization of Interaction Flow Model Feedback.

Model One: Audio Only	Appropriateness of sound types Suitably "safe" choices Simple and straightforward interaction
Model Two: Hierarchy	Countermeasures are interesting Overall model is reversed. The system should become progressively less taxing on a sleepy driver as opposed to more taxing It seems punitive or stigmatizing
Model Three: Sensitivity	How smart is the system? The idea is interesting. Could sensitivity reside with the system, not the user?
Model Four: Modality	The randomness is a concern Appropriateness is an issue Wrong degree of choice Haptic is interesting but chal[illegible]ng

One of the main risks that the experts recognized is the potential for drivers to view the system as a "one more mile in you" device, which they saw as a disturbing position. Rather, they encouraged that the system be positioned as a more friendly, reliable advisor that could be trusted to display accurate information.

Focus Group 2 – CMV Drivers

Several drivers did not describe specific critical incident events, but instead focused on typical drowsiness episodes. Those that had experienced critical incidents were able to recount them in vivid detail, often displaying intense relief that they had survived the episodes.

Critical episodes included rumble strips, hallucinations, unrelated police action, and trance-like periods of degraded awareness. Most stories involved interstate driving and highway speeds. Several drivers described pronounced drowsiness from about 3 a.m. to sunrise. Eye closure periods were usually reported as being three seconds and the word "scary" was frequently used. One driver questioned how far he had driven, while another one was grateful to have received a speeding ticket and believed that the intervention may have saved his life. Drowsiness episodes were reported to have happened on clear nights with empty, straight roads and in inclement weather. In some instances, the drivers described the behavior of fellow CMVs whose driving patterns led them to believe that those drivers were drowsy.

It seemed that all of the drivers understood the gravity of the condition and the consequences that could result while continuing to drive when drowsy, even when circumstances precluded their ability to pull over and rest.

Design Exercise Results

The four teams were in agreement that they wanted brightness adjustment and redundant warnings, i.e., multimodal. The teams reached consensus in the group discussion that the device could be powered by the ignition with the driver having the option to turn the warning system off. Three out of four teams also wanted volume control.

Three out or four teams stated they wanted the most aggressive visual and audio warnings, wanted to be able to select the warning sound, and wanted the warnings to continue until they were acknowledged. One of the teams argued strongly that if a driver has the ability to diminish the warning he or she would do so. In regards to the sounds themselves, mostly aggressive ones were selected and the discussion skewed to aggressive or potentially unsafe sounds, as well as comical ones such as one driver suggesting that his ex-wife's voice be used. The drivers also discussed the possibility of powering down the truck.

On a more serious note, the drivers conceded that the system could be misused to extend hours of service. One driver stated that, "You can only doze so much before you crash." There were also concerns about confidentiality and liability if the device collected and retained information. The discussion ended with general consensus that it "comes down to responsibility and human error" but that this could be a valuable aid for responsible drivers.

DISCUSSION

It became clear from the drivers' comments that there was a slight conflict between how they viewed the device and how the scientific community was positioning the device. The drivers seemed to view the system as a loyal servant that would alert the driver when he became drowsy, while the scientific community viewed the system as a trusted advisor that would encourage the driver to stop and rest. This difference has subtle behavioral implications. The drowsy driver detection and warning system measures drowsiness, not alertness, and as such is not meant to "alert" a driver, i.e., wake he or she up like an alarm clock complete with a snooze alarm. Rather, it is meant to advise the driver that their condition is becoming unsafe and promote safe behavioral choices. Also, the drivers were clearly expressing the need to be in control (servants have masters) and while this is a tenant of usability design, the types of personal control are variable when it comes to a safety critical device.

The final design specification has many features to address both of these views and allow the drivers to retain a degree of control without compromising the warning system.

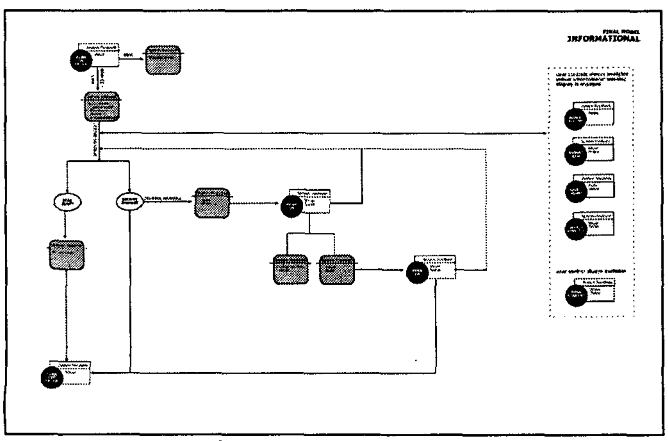

Figure 3: Final interaction flow model

The drivers can control or select many of the features of the interface to correspond to their specific driving environment and individual desires. The interaction model (Figure 3) allows the driver to:

- Adjust the sensitivity of the drowsiness warning to minimize false alarms
- Select sounds that range from a robust alerting sound to a gentle advisory tone
- Adjust the volume to account for the ambient sound environment
- Disable the warning system should he/she find it bothersome
- Dim the functional controls to match their interior running lights

The warning sound occurs simultaneously wi informational warning display. The informational display shows the time the driver's eyes were closed (up to four seconds) and the distance traveled (up to 120 yards, calculated using a speed of 61 mph). A secondary display shows elapsed time between warnings and total warnings received during a drive. Both are preceded by an auditory alert.

Additional detail on the process, interface, and interaction can be found in Ayoob, Grace, and Steinfeld (2002).

Until drivers experience the new design in an operational setting, a gap exists between professed want and actual use, for example, the perception that an aggressive sound is desirable. A functioning prototype is in the process of being piloted with several drivers and will then be used in a large-scale field operational test.

The idea that improved situational awareness can lead to safe decision making needs to be explored further in conjunction with educational efforts for the drivers so that they better understand the physiology behind fatigue and can make off-the-road changes as well.

ACKNOWLEDGEMENTS

This research was supported in part by the National Highway Traffic Safety Administration and Attention Technologies Inc. Special thanks to all of the drivers who drive through the night to support a main infrastructure of our society.

REFERENCES

Ayoob, E., Grace, R., & Steinfeld, A. (2002). *Subtask B, Phase II: Driver-Vehicle-Interface (DVI) Development of a Drowsy Driver Detection and Warning System for Commercial Vehicles, Draft Final Report.* NHTSA Drowsy Driver Detection and Interface Project, DTNH 22-00-D-07007.

Beyer, H., & Holtzblatt, K. (1998). *Contextual Design: Defining Customer-Centered Systems.* San Francisco, CA: Morgan Kaufmann Publishers.

Brown, I.D. (1997). Prospects for technological countermeasures against driver fatigue. *Accident Analysis and Prevention 29,* 525-531.

Dinges, D. F. (1989). The nature of sleepiness: Causes, contexts and consequences. In A. Stunkard and A. Baum (Eds.), *Perspectives in Behavioral Medicine: Eating, Sleeping and Sex.* Hillsdale, NJ: Erlbaum.

Dinges, D. F., Mallis, M., Maislin, G., & Powell, J. W. (1998). *Final Report: Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and as the Basis for Alertness Management* (Report DOT HS 808 762). Washington, D.C.: National Highway Traffic Safety Administration.

Federal Highway Administration. (1998). *Crash Problem Size Assessment: Large Truck Crashes Related Primarily to Driver Fatigue.* Washington, D.C.

Flanagan, J. C. (1954). The critical incident technique, *Psychological Bulletin, 51,* 327-358.

Grace, R., Byrne, V. E., Bierman, D. M., Legrand, J. M., Gricourt, D. J., Davis, R. K., Staszewski, J. J., & Carnahan, B. (1999). A machine vision based drowsy driver detection system for heavy vehicles. *Proceedings of the Ocular Measures of Driver Alertness Conference, April 26-27, 1999,* (FHWA-MC-99-136, pp. 75-86). Washington, D.C.: National Highway Traffic Safety Administration.

Grace, R., & Stewart, S. (2001). The Co-Pilot: A Low-Cost Drowsy Driver and Driver Inattention Monitor. In proceedings of *The International Driving Symposium on Human Factors in Driver Assessment, Training And Vehicle Design, Aspen, CO, August 14-17, 2001.*

Hackos, J, T., & Redish, J. C. (1998). *User and Task Analysis for Interface Design.* New York, NY: John Wiley & Sons, Inc.

Jordan, P. W. (1998). *An Introduction to Usability.* Bristol, PA: Taylor & Francis LTD.

Mallis, M. M., Maislin, G., Konowal, N., Bierman, D. M., Byrne, V. E., Davis, R. K., Grace, R., & Dinges, D. F. (2000). *Biobehavioral Responses to Drowsy Driving Alarms and Alerting Stimuli. Final Report.* (DOT Contract No. DTNH22-93-D-07007). Washington, D.C.: U.S. Department of Transportation.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., & Kirn, C.L. (1994). *Research on Vehicle-Based Driver Status/Performance Monitoring: Development, V[illegible]ion, And Refinement of Algorithms for Detection of Dr[illegible] Drowsiness. Final Report.* (DOT HS 808 247). W[illegible]gton, D.C.: National Highway Traffic Safety Administration.

Wierwille, W. W., Lewin, L.G., & Fairbanks, J, R. (1995). *Research on Vehicle-Based Driver Status/Performance Monitoring: Seventh Semi-Annual Report.* (DOT HS 808 299). Washington, D.C.: U.S. Department of Transportation.

Wylie, C.D., Shultz, T., Miller, J.C., Mitler, M.M., & Mackie, R.R. (1996). *Commercial Motor Vehicle Driver Fatigue and Alertness Study: Project Report.* (Report No. FHWA-MC-97-002). Washington, D.C.: U.S. Department of Transportation.