UNDERSTANDING DRIVER - AUTOMATED VEHICLE INTERACTIONS THROUGH WIZARD OF OZ DESIGN IMPROVISATION

Brian Ka-Jun Mok, David Sirkin, Srinath Sibi, David Bryan Miller, Wendy Ju
Stanford University
Stanford, CA, USA
{brianmok, sirkin, ssibi, davebmiller, wendyju}@stanford.edu

Summary: This paper describes a Wizard of Oz study that was performed to gather insights on how automated vehicles (NHTSA's Levels of Automation 2 and 3) should interact with human drivers. Twelve design improvisation sessions were conducted inside a driving simulator with interaction and interface design experts. The participants drove through a simulated course with various terrain and road conditions, while the two human operators (wizards) controlled the audio and driving behavior of the car. Through the feedback collected in these sessions, insights in five areas were discovered: drivers' desire for shared control, transitions in driving mode, response latency, addressing requests, and drivers' trust in the car. Additional examining yielded potential concepts and ideas that may be implemented and tested in future work.

INTRODUCTION AND OBJECTIVES

With the advent of partially and fully automated vehicles, drivers are increasingly sharing control of driving with their cars. How should these automated vehicles interact with human drivers? If the driver and the vehicle are considered to be a team collaborating to accomplish a driving task (Inagaki 2009), it is important that they both understand each other. If the car is unable to interact properly, this could lead to lower levels of trust and comfort for drivers. Therefore, it is critical to understand what types of messages should be sent and what behaviors a car should perform in order to effectively communicate and interact with the driver. This paper presents a Wizard of Oz study that was conducted to better understand the questions and concerns held by drivers of automated cars, and to identify key issues in transitioning and sharing control. With the help and feedback of interaction and interface design experts, we conducted twelve design improvisation sessions with our Wizard of Oz autonomous driving simulator system to determine important aspects governing the driving experience in an automated vehicle.

BACKGROUND AND RELATED WORK

In Wizard of Oz studies, participants are told to act as if they are interacting with a computer system through an interface, when in fact their interactions are mediated by a human operator - the wizard. The use of "the wizard in the loop" experimental set up allows experiments to be both less constrained - through use of improvisation or the wizard's expressiveness and more systematically constrained - by cutting out the limitations of an automated system - than would be possible with a real computer-operated system (Dahlbäck 1993). Wizard of Oz can provide an inexpensive and flexible simulation due to the use of a human operator (Cross 1977). This technique can be used for testing systems, or also as an iterative design methodology. As indicated in Hoffman et al. (2014), Wizard of Oz experiments early in the design process can be used to explore a wide range of possibilities and can help identify key features and functions of a system. Given its versatility, Wizard of Oz is a good platform to examine interactions between humans and automated cars. The Wizard of Oz technique is commonly used in the automotive research community for the design and study of automotive user interfaces. It has been used by the designers of Virtual Intelligent Co-Driver (VICO) to evaluate user expectations (Geutner 2002), by developers of speech-based in-car entertainment systems, and by researchers at TU Munich (Schuller 2006). This technique has also been used by researchers developing gesture-based interfaces for secondary tasks in a car environment (Alpern 2003), by researchers looking at different automotive interface input modalities (Tsihoni 2004) and by researchers at Volkswagon ERL to examine speech in cars (Lanthrop 2004).

METHODOLOGY

The Simulator

The Driving Simulator (Fig. 1) is an immersive full-car simulator. The system consists of two parts: the display and the car. The display is a 270-degree field of view screen. This cylindrical display, centered on the driver, utilizes five projectors to display the simulated environment. Software stitches the displays of these five projectors to create one large seamless display. Another projector is used to display the rear view. We utilize a modified Toyota Avalon to provide participants with a realistic interface. A motor attached to steering wheel and a pneumatic system powering the pedals provides realistic force feedback to the participants. LCD panels are installed in the side view mirrors to complete the immersive experience.

**Figure 1: The Driving Simulator is used for experiments on Automated Driving.
It features a Toyota Avalon chassis and a 270-degree frontal arc seamless curved screen.**

The Wizard Station

The Wizard of Oz station (Fig. 2) is designed to be operated by two human "wizards." The Interaction Wizard is tasked with observing and interacting with the participant, while the Driving Wizard mainly interacts with and manipulates the simulation environment. The wizards can both hear the participant through the use of a small microphone hidden within the car. The wizards are able to observe the participants and their context in real-time. Three GoPro cameras strategically placed to capture / observe the participant's face, the broader driving cabin, as well as the driver's footwell. In addition, the wizards have a feed of the front view from the simulator, as well as the dash and the center console display. The camera views and the simulator feed are simultaneously recorded.

In this experimental setup, the Interaction Wizard “talks” to the participant as the voice of the automobile, utilizing a text to speech reader software called Ivona. We recorded the entire audio conversation between the car and the participant. Out of the many voices that the software provided, we decided to use Salli, a “Female, American-English Voice” for this study because it was the most realistic and natural sounding voice. The audio from the software is transmitted to a Bluetooth receiver linked to the car’s speaker system. Acting as the car, the wizard is able to listen and respond to any verbal requests or questions that the participant has. This wizard is also given some control over the occurrence of the events in the simulated course through use of buttons on the dashboard that can trigger events and conditions in the simulator.

Figure 2: Wizard of Oz Station. Wizard 1 monitors and interacts with participant, Wizard 2 interacts with and manipulates simulation environment.

The Driving Wizard is also responsible for the car’s driving during automated mode. A set of buttons on the dashboard will transfer control from the participant in the car to the Driving Wizard, who can then steer the car using a force feedback driving controller. The wizard can steer, accelerate, brake, and toggle the turn signals. Maneuvers from this steering wheel are fed back into the simulator steering wheel so that the participant’s steering wheel turns in accordance with the actions taken by the Driving Wizard on behalf of the automated vehicle. A screen outputting the center display allows the wizard to effectively navigate through the course in automated mode. Also, control can be restored to the participant in the car at any time.

The Study

To understand how the automated vehicle should interact, we conducted a study inside the driving simulator and invited twelve interaction and interface design experts to act as the participants. As these were design improvisation sessions, the participants were given more freedom to interact as they wanted. They were not provided with any instructional paradigms about the car and were allowed to make requests / ask questions at any time. Wizards interacted in real-time with the participant through the car’s interface throughout the whole course. The driving course is composed of four parts: a training section and three event sections which span a variety of possible driving contexts where expectations and behaviors around automation may be different. The training section is intended to help participants familiarize themselves with driving in the simulator. This section contains intersections, roundabouts, and various other road terrain features. The first event section contains forests and hills. The second event section contains a city and involves urban driving. The last event section contains a high speed freeway. These sections contain critical events such as road hazards that the wizard can activate on demand.

Our study protocol focused on changing the level of autonomy and control throughout the course, as well as the interplay between pushing or pulling information from the drivers during interaction. The car would offer to take over driving at certain points throughout the course, and it would also request the participant to take over control at other times. The protocol around the pushing and pulling of information between the participant and the vehicle was intended to follow natural conversation norms, so this aspect of the study was not controlled. The wizards would be able to either freely offer explanations for the car's automated driving (pushing information to the driver), or respond to queries about the car's automated driving behavior (pulling of information by the driver). After the study, participants were interviewed to gather additional insights. They were asked a set of standard questions such as "At what point did you start to trust the car?" Participants were also asked questions related to their individual experiences such as "Why did you try to instruct the car instead of taking over control?" These questions lead to the discovery of some results and observations about key aspects of the driving experience and desired interaction patterns.

RESULTS, OBSERVATIONS AND DISCUSSION

Desire for Shared Control

In the course of the drive, there were segments when the car intentionally drove imperfectly during automated mode. The car either drifted laterally slightly or went too close to objects in front of it. According to the feedback provided by the participants, those moments caused the participants to want more control of the car. Although the driving was slightly flawed, the participants still had the trust/confidence in the car's performance and did not want to completely disable the automation. They only wanted to provide inputs to correct the car's driving behavior, with the car staying in automated mode the entire time. All drivers at some point had tried to turn the steering wheel or step on the gas/brake pedals, but did not ask for the control back. In this desired automated driving mode, the car would perform most of the driving tasks and the drivers would only provide corrective assistance during times of concern.We note that this form of drivers' desire for shared control is not classified directly under National Highway Traffic Safety Administration's Current Levels of Automation Model. For this mode desired by participants, the driver acts more as an overseer who directs the overall driving strategy rather than an operator who provides constant input to the car. It can be seen as the inverse of NHTSA's Level 1: Function Specific Automation, where the driver is still mainly performing the driving task and the car intervenes (NHTSA 2013). This mode is also different from NHTSA Level 2: Combined Function Automation and Level 3: Limited Self-Driving Automation, where the car performs most or all of the driving tasks, and the driver takes complete control only when a problem occurs. This suggests that there may be other intermediate levels of autonomy that will provide an alternative path to Level 4: Full Automation.

However, in what ways should this level of automation function? Many of the participants wanted the car to change its driving behavior when they turned the steering wheel or when they pressed the gas/brake pedals. They then expected the car to continue in automated mode after the input had been received. This can be challenging, as there may be uncertainty around when the driver decided to stop providing instructions to the car. Furthermore, some participants tended to keep their hands on the steering wheel to be aware of what the car was doing. Others did that as a method to deal with discomfort. Considering these tendencies, it is difficult to determine the driver's intent. Is the driver trying to influence the car, trying to observe the car, or performing subconscious actions? An alternative method that participants attempted to use was to give the car specific audio cue. Participants said phrases such as "You are drifting to the right" or "You are getting too close to the car in front of us" to indicate the desired corrections. Although the use of audio may be less structured, the intent of the driver appears to be clearer in those cases.

Transitions in Driving Mode

Although we utilized both audio and visual indicators, our participants found that transitions between manual and automated driving modes were often unclear. In particular, it was difficult to determine the exact moment of transition using speech. Even when the car used the phrase "I have control now," with a visual indicator on the dashboard changing afterwards, participants were still uncertain when they should actually relinquish control. After the transition, many of the participants still asked, "Are you driving the car now?" and, "Can I let go of the steering wheel now?" We need to be able to communicate the instant when change occurs. Therefore, sentences and phrases alone are not sufficient as the message is conveyed over a period of time. Adding a short chime led participants to be much clearer about the demarcation of the transition. The chime was a familiar alert to the participants, which when added to speech would provide an adequate notification and explanation of what was occurring. A count down was also effective; however, participants felt it took too long and was annoying.

In addition to the audio indicators, we also examined different graphical methods to indicate that a transition was going to occur. In the instrument cluster, we placed a mode indicator graphic. When the car was in manual mode, the graphic would be gray and displayed "Autonomy Off." Similarly, when the car was in automated mode, the graphic would turn green and display "Autonomy On." Participants found this visual cue to be useful to confirm that a transition had occurred, but often did not notice right when it changed. Citing their uncertainty around the moment of transition, some of the participants suggested the use of haptic feedback devices could also provide an adequate notification. Having the steering wheel vibrate, the seatbelt tighten, or the seat move could all be potentially effective ways to alert the driver of this change.

Response Latency

All the participants found the time it took for the car to respond to questions and requests to be acceptable, which indicated that this Wizard of Oz setup could be operated at the appropriate timescale. Limited by the time the wizard needed to type the message, most of the replies from the car were given within 10 seconds. This limitation was particularly apparent when a complex answer needed to be given, as an uncomfortable period of silence was created. This delay often caused participants to question whether the car was able to interpret what they said.

One of the interaction methods that appeared to ameliorate the latency problem was to use an acknowledgement before creating a more detailed response. For example, using the phrase, "Let me find out," immediately after the participant asked a question would let the participants know that their statements were acknowledged. Also, participants felt that they were still engaged with the car, allowing the wizard a longer period of time to respond to the participant before the interaction became uncomfortable. Another suggested solution by the expert participants was to use sounds that indicated the car was processing the information. However, there were still limitations with the above methods. Often the expected response time was contextual. For instance, some of the participants wanted information on an upcoming building, they expected a timely response, before the car passed the building. One participant noted, "Maybe I would have liked to modify my trip, but now it is too late." So, there are still situations where the delay inherent to both to speech recognition will be difficult to mitigate.

Addressing Requests

Participants often instructed the car to do certain tasks such as, "pass that slow vehicle in front of us" or, "tell me about today's news headlines." Occasionally, these requests could not or should not be performed by the car. However, it would be extremely important to find out the car's reasoning why certain requests could not be completed. If the car presented a technical reason why it could not do certain jobs, the participants were normally quite satisfied with the response and did not argue further. An example of this was when a participant asked to "play music on his play list." The car responded that, "I do not have access to the files and could not complete the request." The participant acknowledged this limitation and continued to converse with the car without becoming adversarial. On the other hand, when the participants knew that the car was capable of performing the request but refused to do so due to non-technical reasons, they were less willing to relent. One example of this was the participant asked the car to "go faster" but the car would not because "the speed limit was 35 miles per hour." In this case, the participants continued to ask the car to speed up. Several participants even chose to ignore these safety reminders, opting to disengage automation and drive significantly faster than the limit.

Trust in Autonomy

There were several common actions performed by the car that helped build the participants' trust and confidence in the automated mode. One of these was the car being able to traverse through difficult sections of road perfectly. For many participants, after the car was able to stably drive through a sharp S curve, trust and confidence in the automated mode increased. Another action that promoted trust with participants was the car's situation awareness. One participant cited that the car pointing out curves and hills up ahead made the car seemed more secure/reliable. Likewise, another participant noted that the car mentioned being able to see the pedestrians in the environment was reassuring. The participants also indicated that they had trust the car more if it was able to successfully avoid crashes. One of our participants tried to "test" the automated system by not providing any input. As the car was still able to avoid a cutoff car on its own, the participant felt that it was okay to relinquish control. It was also observed that when the car drove imperfectly, participants were more alert and vigilant, not fully trusting the automation. Drivers' trust and confidence in the car could also be repaired over time. After experiencing imperfect driving by the car, one participant immediately disengaged automation. Attempts to reengage automation shortly after were all met with disapproval. However, after interacting for another 15 minutes, the car was given another opportunity to use automated mode by the driver.

CONCLUSION

Through conducting the twelve design improvisation sessions with our interaction experts, we are able to find some important insights concerning automated driving. Participants wanted to share control with the car without taking over full control. They would like to know exactly when a mode switch happen and need to be alerted by the car in a clear manner. Delays in response and unperformed requests were acceptable as long as the car provided the participants with proper responses. Finally, the car had a multitude of methods to help build trust with the participants. Additional concepts and ideas, such as different methods to provide alerts, are generated through this process. We intend to implement and test them in future work.

ACKNOWLEDGEMENTS

The research project is supported by Toyota CSRC. The authors thank James Foley, Mishel Johns, Barbara Karanian, Neeraj Sonalker, Karen Robinson, William Verplank, Jamy Li, Matt Kresse, Karlin Bark, Nuri Kim, Michael Sturtz, David Ayman, Jofish Kaye, and Erin Panttaja for their contribution to this study.

REFERENCES

Alpern, M., & Minardo, K. (2003, April). Developing a car gesture interface for use as a secondary task. *CHI'03 extended abstracts on Human factors in computing systems* (pp. 932-933). ACM.

Cross, N (1977). *The Automated Architect*. Pion Limited. p. 107.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies—why and how. *Knowledge-based systems*, *6*(4), 258-266.

Geutner, P., Steffens, F., & Manstetten, D. (2002). Design of the VICO Spoken Dialogue System: Evaluation of User Expectations by Wizard-of-Oz Experiments. *LREC*.

Hoffman, G., & Ju, W. (2014). Designing Robots With Movement in Mind. *Journal of Human-Robot Interaction*, *3*(1), 89-122.

Inagaki, T. (2009). Human-machine collaboration for safety and comfort. *Proceedings of ENRI International Workshop on ATM/CNS*. Tokyo, Japan.

Lathrop, B., Cheng, H., Weng, F., Mishra, R., Chen, J., Bratt, H (2004). A Wizard of Oz framework for collecting spoken human computer dialogs: An experiment procedure for the design and testing of natural language in-vehicle technology systems. *Proc ITS*

National Highway Traffic Safety Administration. (2013). *Preliminary Statement of Policy Concerning Automated Vehicles*. Washington, DC: Department of Transportation.

Schuller, B. Lang, M., & Rigoll, G. (2006). Recognition of spontaneous emotions by speech within automotive environment. *FORTSCHRITTE DER AKUSTIK*, *32*(1), 57.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech recognition versus a touch-screen keyboard. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46* (4), 600-610.