HUMAN FACTORS EVALUATION OF LEVEL 2 AND LEVEL 3 AUTOMATED DRIVING CONCEPTS

Paul, Rau
National Highway Traffic Safety Administration
United States

Myra, Blanco
Jon, Atwood
Holland, Vasquez
Tammy, Trimble
Vikki, Fitchett
Josh, Radlbeck
Gregory, Fitch
Sheldon, Russell
Virginia Tech Transportation Institute
United States

Paper Number 15-0428

ABSTRACT

This project evaluates how drivers interact with different automated vehicle functions under various concepts of Level 2 and Level 3 automation. The objective is to determine whether principles for human-machine interface (HMI) design for automated vehicles could be based on things such as timing, sequence, and presentation of automated functions produced by this study. Methods involve test track evaluations of participants using three distinct automation concepts, two involving automation Level 2 and one involving automation Level 3 (as defined by the National Highway Traffic Safety Administration [NHTSA] policy paper on vehicle automation; NHTSA, 2013). Data sources included both objective and subjective data from participants’ responses to the different portions of the experimental protocols. Results will be produced from parametric linear regression analyses and qualitative evaluations of participants’ subjective responses to questionnaires. Where appropriate, statistical techniques will be applied for conditioning the sample data to ensure that the assumptions underlying these analyses are met. The detailed timing, sequence, and presentation measurements from the various research efforts involved herein will be used to specify human factors design principles for automated vehicle HMIs. The resulting principles would benefit from subsequent naturalistic evaluations for fine-tuning the performance metrics, and for addressing any gaps or new questions arising from this research. Crash avoidance technologies are evolving rapidly toward increasing automation, involving a higher complexity of interoperability between user and vehicle functions than what has previously been known. Understanding the detailed human factors capabilities and limitations of these users and the impacts of the timing, sequence, and presentation of information presented to the users will be important for shaping the safety policies.

INTRODUCTION

Technological advancements over the past decade have led to the emergence of advanced driver assistance systems. Current features such as Adaptive Cruise Control (ACC), collision warning, and automatic braking systems are becoming commonplace in modern automobiles. Furthermore, automated systems that combine limited lateral and longitudinal control over a vehicle are becoming commercially available. Some of these systems incorporate various methods to ensure driver participation.

While automated systems offer the potential for increased safety and reduced human error, their use may create issues which could benefit from further investigation. These issues may include negative adaptations based on misunderstanding, misuse, over-reliance on the automated systems, and distractions from the driving task due to interaction with the automated system. These issues should be examined in order to address any potential for unforeseen consequences of increased automation.
Of specific interest is how an automated system will impact operators’ willingness to engage in non-driving-related tasks. As noted in previous work, (e.g., Llaneras, Salinger, & Green, 2013), the current generation of automated systems is designed to support, rather than replace, the driver. The presence of automated systems may create the perception to free an operator’s attention, which may then be directed at non-driving tasks. The redirection of the operator’s attention to non-driving tasks may also impact an operator’s situational awareness, including the ability to perceive critical factors in the environment or to detect issues with the automated systems (e.g., system state changes or failures).

As automated driving technology advances, the “driver’s” role is shifting from active vehicle control to passive supervision of the automated system and/or the environment. The current study focused on the human factors issues that arise when vehicles equipped with automation technologies shift the human from the role of driver to that of operator. Automated vehicle systems must be designed to instruct and prompt the operator to act, if and when needed, in a timely and appropriate way in order to ensure safety. This study investigated how operators interact with partial automation under National Highway Traffic Safety Administration (NHTSA) Levels 2 and 3 (NHTSA, 2013). Level 2 (Combined Function Automation) and Level 3 (Limited Self-Driving Automation) were of particular interest because this is the point at which the driver’s role transitions from one of driving to one of operation.

Objective

The purpose of this study was to examine the interaction between human users and automated vehicle systems. Specifically: how do human users interact with vehicles that have L2 and L3 automated systems, can these users take over control of the driving task when required, and can they determine the acceptable balance between controlling the vehicle when necessary and letting the automated system function as designed to perform the driving task when appropriate? The ultimate goal of this research was to ascertain how operators interact with automated vehicles and determine how automated vehicle technology can best support safe driving.

Project Research Questions

This study centered on six key research questions developed by NHTSA. The focus of this research was to address each question based on sound empirical research findings. The research questions were:

1. How do drivers interact with and operate vehicles that offer Level 2 and Level 3 automation; e.g., what is the driver performance profile over length of time in continuous or sustained automation?
2. What are the system performance risks from driver involvement with, and interruption from, secondary tasks (such as portable electronic device use) that could arise when operating Level 2 or Level 3 automated vehicle systems?
3. What are the most effective hand-off strategies between the system and the driver, including response to faults/failures?
4. How do drivers engage, disengage, and reengage with the driving task in response to the various states of Level 2 and Level 3 automation?
5. How do drivers perform under various operational concepts within Level 2 and Level 3 automation, such as systems intended for everyday driving on open roadways in mixed traffic or systems intended for dedicated roadway-vehicle applications (e.g., automated lanes, remote highways)?
6. What are the most effective human-machine interface concepts, guided by human factors best practices, which optimize the safe operation of Level 2 and Level 3 systems?

The six aforementioned research questions were addressed in three experiments. Experiment 1 examined how best to alert operators to regain control of the vehicle, Experiment 2 examined the system prompt effectiveness (i.e., how effectively the HMI communicated to the operator) over time, and Experiment 3 examined operator behavior over time. Experiments 1 and 2 were conducted using a vehicle equipped with an L2 system, while Experiment 3 was conducted with a vehicle equipped with an L2 system that can simulate L3 driving on a test track. The details of the experiments and their expected findings are presented in the following sections.
EXPERIMENT 1: ALERTING OPERATORS TO REGAIN CONTROL OF AN L2 AUTOMATED VEHICLE

Purpose

There are numerous ways to notify operators that they need to regain control of a partially automated vehicle. The purpose of this experiment was to investigate which human-machine interface (HMI) characteristics are most effective at issuing a Take-Over Request (TOR) and to identify the transition times between the operator and the automated functions in regard to the driving tasks.

Participants

Data were collected from 35 participants recruited from the greater Detroit, Michigan area; however, 10 participants were considered invalid (i.e., session cancellation due to adverse weather, track closures, or technical issues associated with the prototype vehicle). As such, the analysis will represent data from 25 participants (16 males, 9 females). The mean age of participants was 44.3 years old (S.D. = 19.24), with ages ranging from 18 to 72 years old.

Method

A single, long-exposure, experiment was conducted. Participants were provided with a thorough familiarization of the vehicle and its operation, followed by a single, approximately 90-minute, exposure to the vehicle in L2 automated driving. During the driving session, participants were instructed to perform non-driving-related tasks (e.g., e-mail, web browsing) and were, at times, presented with alerts stating that they must take control of the vehicle. Three forms of alerts were presented: Cautionary, Imminent, and Staged. The Cautionary alerts provided information to the participants that a potential problem was detected. The Imminent alerts provided the participants with a message that an active fault was detected. The Staged alerts transitioned from a cautionary alert phase to an imminent alert phase. Participants’ reactions to these messages, both in duration and method of response, were among the variables examined in this experiment.

Experimental Design

The study was performed as a within-subject design. All participants completed one 90-minute driving session during which they received a total of 19 system alerts. For each of the alert types, participants experienced three unimodal alerts (visual only) and three multimodal alerts (visual + haptic). The study was designed to mimic worst-case scenarios when the operator was not monitoring the roadway due to the non-driving tasks.

The alerts were presented to participants in six different orders. Each order consisted of the six different combinations of alert type and alert modality, and this was repeated three times within the experimental session. Using all six possible alert type and alert modality combinations, a Latin square was developed to create six different orders of alert presentation. The order was repeated three times within the 90-minute driving session, resulting in 18 alerts. After receiving these 18 alerts, each participant received an Imminent multimodal alert coupled with an experimenter-triggered lane drift, resulting in a total of 19 alerts. The alerts were presented at random times between 2 and 8 minutes; thus, participants were less likely to be able to anticipate when they would occur.

Venue

This experiment was conducted at the Milford Proving Ground circle track in Milford, Michigan. This facility is owned and maintained by General Motors (GM) and includes a 4.5-mile banked circle track with five travel lanes. The travel speed for each lane falls within a designated speed range, with the innermost lane allowing for stop and go traffic and the outermost lane being restricted to speeds of 100 mph and above. Experiment 1 was conducted in Lane 3, which allowed speeds of between 50 - 70 mph.
Vehicle

A 2009 Chevrolet Malibu equipped with a prototype L2 automated driving system was used (See Figure 1). As part of the automated driving system, several HMI components were installed and the vehicle was modified to include ACC and lane centering along with a flexible driver interface. Additionally, the vehicle was equipped with a researcher’s control console, which was designed to allow the in-vehicle experimenter to trigger various displays and to change the operation of the automation systems, including simulating erroneous behavior and equipment failures.

![2009 Chevrolet Malibu with a prototype L2 system used in Experiment 1.](image)

The vehicle was also equipped with a data collection and recording device. These key variables were collected: status of the automation (e.g., off, on and actively controlling, failure mode), vehicle speed, lane position, and flags indicating the presentation of messages and system failures. In addition, video views which included the operator’s face, forward roadway, and HMI were collected.

Study Session

Prior to beginning the study session, participants were provided with a static orientation to the experimental vehicle, which included information about the basic controls and the L2 automation features. Following this, participants received an on-track orientation consisting of four laps on the test track. During the 90-minute driving session, participants were instructed to perform a variety of non-driving tasks using a tablet computer when the L2 automation was activated. Participants were presented with three types of non-driving tasks to complete using the tablet computer: navigation, email, and web-browsing. At approximately 5-minute intervals (in random values ranging from 2 to 8 minutes), participants were provided with unimodal (visual only) or multimodal (visual + haptic seat vibration) alerts (Cautionary, Imminent, or Staged) instructing them to take control of the vehicle. Details pertaining to the Cautionary, Imminent, and Staged alert timings are depicted in Figure 2.
Figure 2. Cautionary, Imminent, and Staged take control alert timelines for Experiment 1.

A trust scale was administered 10 times throughout the experimental session at approximately 9-minute intervals. Participants were asked to rate their trust in the ability of the automation to function properly while they engaged in non-driving tasks using a 7-point Likert-type scale. In addition to the 10 trust ratings collected throughout the experimental session, participants were asked to complete the after-experience trust scales and participate in an open-ended interview upon completing the driving session. Compensation was provided for participation in the study.

EXPERIMENT 2: SYSTEM PROMPT EFFECTIVENESS OVER TIME

Purpose

The second experiment investigated how to prompt operators to monitor the driving environment when engaged in a non-driving-related task during the operation of an L2 automated vehicle. A secondary purpose was to investigate the effectiveness of the prompts over time.

Participants

Data were collected from 56 participants recruited from the greater Detroit area (28 males, 28 females) with a mean age of 41 years old (S.D. = 16.3), with ages ranging from 18 to 72 years old.
Method

A single, long-exposure experiment was conducted. Participants were provided with a brief familiarization with the vehicle and its operation, followed by three 60-minute experimental sessions. During the sessions, participants were given tasks to be completed using a tablet computer. During these tasks, participants received prompts based on their predetermined prompt condition (either 2-second, 7-second, or No Prompts). For the 2- or 7-second prompt conditions, participants received prompts after periods of inattention to the driving environment for the corresponding amount of time. Participants given the No Prompts condition did not receive any prompts and they were free to behave as they thought was appropriate.

In addition to these prompts, at a random time during one predetermined session, the participant received an alert for a surprise left lane drift, consisting of a haptic seat alert and a flashing red LED. In a different predetermined session, the participant experienced a surprise lane drift with no alert, which consisted only of a left lane drift without any alert and with the prompting system disabled. The experimenter-injected lane drift was used to simulate a lane-keeping performance issue combined with a failure of the prompting system. Note that, to the participants with the 2-second and 7-second prompt conditions, the alert that they received along with the lane drift was indistinguishable from the prompts that they had been receiving based on their attention state. Participants’ reactions to these prompts, alerts, and lane drifts, both in duration and method of response, were examined in this experiment.

Experimental Design

The study was performed as a 3 x 3 x 3 mixed factorial design. Each participant completed three successive driving sessions, and each session included one of the following: a lane drift with an alert, a lane drift without an alert, or no lane drift. Participants experienced each of these conditions once during the experiment. In addition, there were also three different prompt conditions that were used with the driver monitoring system, and each participant experienced only one prompt condition, either: 2-second, 7-second, or No Prompts. The prompt timing was based on previous distraction research (2-second prompts) (e.g., Klauer et al., 2006) and expert opinion (7-second prompts). Additionally, the study was designed to mimic worst-case scenarios when the operator was not monitoring the roadway due to the non-driving tasks.

Venue

As was the case for Experiment 1, Experiment 2 was also conducted at GM’s Milford Proving Ground circle track in Milford, Michigan. However, this experiment utilized Lane 2, which allowed speeds of between 30-50 mph.

Vehicle

A 2010 model year Cadillac SRX equipped with a prototype L2 automated driving system was used as the experimental vehicle (See Figure 3). As part of the automated driving system, several HMI components were installed. These included an instrument panel binacle-mounted screen providing information on the automated driving system, and two steering wheel buttons to control the automation: one ACC button, and one button for the lane-centering system, a prototype automated vehicle system.
The vehicle was equipped with Virginia Tech Transportation Institute’s (VTTI) data acquisition system (DAS). The variables collected by the DAS included status of the automation, vehicle speed, and lane position. In addition, video views including the operator’s face, forward roadway, and HMI were collected.

Study Session

Prior to beginning the study session, participants were provided with a static orientation to the experimental vehicle, which included the basic controls and the L2 automation features. Following this, participants received an on-track orientation consisting of four laps on the test track. Participants then completed three driving sessions, with each lasting approximately 60 minutes. Participants were instructed to begin interacting with a variety of non-driving tasks during the driving session upon activating the L2 automation. Participants were presented with three types of non-driving tasks: navigation, email, and web-browsing. These tasks were similar in terms of the visual/manual demand required and they were presented in a random order.

Each participant was assigned a prompt condition: either 2-second, 7-second, or No Prompts. The driver monitoring system provided three stages of prompts based on the assigned prompt condition and the participant’s attention state. If the participant’s attention state was not on the driving environment, the system provided alerts based on the assigned prompt condition. For the 2-second prompt condition, the prompts began after the participant’s attention state was not on the driving environment for 2 s. For the 7-second prompt condition, the prompts began after the participant’s attention state was not on the driving environment for 7 s. Participants who were assigned to the No Prompts condition did not receive any prompts. The driver monitoring system provided three progressive stages of alerts. Details pertaining to the prompt stages are detailed in Figure 4 below.
During the experiment, each participant experienced both types of lane drift (with and without an alert) and no lane drift—one time each—in different driving sessions, and at random times. All of the surprise lane drifts were prescribed and injected into the condition of interest using the experimenter console. The lane drifts with the alerts represent the condition of a lane-keeping performance issue in which the system warns the vehicle operator in order for him/her to regain control. The situations with no alerts represent conditions where there is a lane-keeping performance issue and a simultaneous failure of the prompt system, but the system does not warn the vehicle operator. Participant responses to the attention state prompts and experimenter-injected lane drifts were measured using visual evidence from the DAS. After completing three driving sessions, participants were instructed to exit the circle track and return to the preparation area. Participants were then interviewed, asked to complete the after-experience trust scales and interview, and provided with a debriefing as to the purpose of the study. Compensation was provided for participation in the study.

EXPERIMENT 3: HUMAN-AUTOMATION SYSTEM PERFORMANCE OVER TIME

The purpose of the third experiment was to investigate what HMI characteristics are effective at alerting operators to regain control of an L3 automated vehicle and to identify the transition times between the operator and the automated functions in regard to the driving tasks.

Participants

Data were collected from 37 participants recruited from the greater Roanoke, Virginia area; however, 12 participants were considered invalid (i.e., session cancellation due to adverse weather or technical issues associated with the DAS or the prototype vehicle). The analysis will consist of data from 25 participants. The mean age of the participants was 38.8 years old (S.D. = 13.77), with ages ranging from 18 to 69 years old.

Method

A single-exposure experiment was performed. Participants were provided with a thorough familiarization with the vehicle and its operation, including use of the automated features. Training was followed by three 30-minute experimental sessions. During the sessions, participants had free exposure to a non-driving-related task (i.e., use of their own cell phone or the provided tablet, as they felt was appropriate) and were presented with a message stating that they must take control of the vehicle. Participants’ reactions to these messages, both in duration and method of response, were examined in this experiment.
Experimental Design

The study was performed as a within-subject design. Each participant completed three successive driving sessions, each session with one of three alert types; all participants received all alert types exactly once. The three alert types were: Staged, Imminent–No External Threat, and Imminent–External Threat. They received a Staged alert in the absence of an external threat, an Imminent alert in the absence of an external threat, and an Imminent alert in response to an external threat (i.e., a revealed box on the road). The Staged alert was composed of four phases: 1) a short tone followed by an informational message asking operators to prepare for manual control (including a countdown timer), 2) a Cautionary verbal alert played in addition to an animated HMI display with the instruction to “please turn off autodrive” presented for 10 s, 3) a repeated cautionary tone played in addition to an orange visual alert stating to “turn off autodrive now” presented for 10 s, and 4) a repeated imminent tone played in addition to a red visual alert stating to “turn off autodrive now” presented for 10 s combined with the automation beginning to apply the brakes. The Imminent alert was composed of a red visual alert stating to “turn off autodrive now” presented for 10 s along with the automation applying the brakes.

Venue

This experiment was conducted on the Virginia Smart Road test track, which is located at VTTI in Blacksburg, Virginia. The test track is constructed to state and federal roadway standards and has a length of 2.2 mi, with looped turns at either end. The straight section of the track is approximately 0.5 mi in length. Two lanes run the duration of the track, with the exception of the looped turns. Wireless Internet coverage is available on the track. The facility is closed to outside traffic and only study-related vehicles were present during the experiment.

Vehicle

A 2012 Lexus RX450h was used as the experimental vehicle for Experiment 3 (See Figure 5). This L2 vehicle was equipped with a prototype automated driving system that can simulate L3 driving on a test track. As part of the prototype system, several HMI components were installed. These included an instrument panel binnacle-mounted screen providing information on the automated driving system, and two steering wheel buttons to control the automation: one ON button on the left side of the wheel and one OFF button on the right side of the wheel.

![Figure 5. 2012 Lexus RX450h with a prototype L3 system used in Experiment 3.](image)

The vehicle was instrumented with VTTI’s DAS. The variables collected by the DAS included throttle/brake input and automation state. In addition, video views including the operator’s face, the forward roadway, and the HMI were collected.

Study Session

Prior to receiving any hands-on training, the participants viewed a 10-minute video summarizing the vehicle’s features with a specific focus on the automated components and operation of the vehicle. This video was a training requirement of the automated vehicle provider and was consistent with the recommendation in NHTSA’s Preliminary Statement of Policy Concerning Automated Vehicles in the section entitled “Recommendations Concerning State Activities Related to Self-Driving Vehicles” (NHTSA, 2013; pp. 10-11). This video was intended to detail the prototype system’s operating capabilities and limitations. The participants were also shown the different
types of the alerts (i.e., Staged, Imminent–External Threat, and Imminent–No External Threat) during the video. Following this, participants received an on-track orientation consisting of four laps on the test track.

The experiment consisted of three 30-minute driving sessions. Participants were able to freely engage in non-driving tasks (i.e. tablet computer and cell phone use) when the automated system was activated. At a randomly selected point within each session, one of the three alert types was presented: Staged, Imminent–External Threat, or Imminent–No External Threat. The alerts happened at a predetermined location and participants experienced all three alert types during the experimental session. While each participant experienced all of the alert types, they were not always experienced in the same order. Details pertaining to Staged and Imminent alert timing and presentation are shown in Figure 6.

![Figure 6. Staged and Imminent take control alert timelines for Experiment 3.](image)

A 15-minute break was offered after each session to allow for participant comfort; however, some participants chose to forgo the breaks. The maximum speed for all sessions was 45 mph, with lower speeds used for the turns at both ends of the track.

The trust scale was presented at 10-minute intervals during each session (at the beginning of the session, followed by administrations after 10, 20, and 30 minutes). Upon completion of the third session, the participant was instructed to deactivate the vehicle’s automation, assume manual control over the vehicle, exit the track, and return to the preparation area. An interview was performed at that point. Compensation was provided for participation in the study.

**PLANNED STATISTICAL ANALYSES**

Because the data are measured repeatedly on participants over time, longitudinal statistical methods—used widely in experiments with repeated measures—can be used to analyze the data. For the variables that involve a time until the participant performs some action (such as regaining control of the vehicle), continuous data methods can be used, although the presence of extreme values may require the use of a data transformation (such as the logarithmic transformation) to effectively normalize the data to help fulfill the assumptions of these techniques. For other types of variables, such as the number of non-driving-related glances (which are counts) and the monitoring rate (a proportion), more general longitudinal methods can be used that account for the wide variety of distributions that these variables follow.

**EXPECTED RESULTS**

The detailed timing, sequence, and presentation measurements from the various research efforts involved herein could be used to develop human factors design principles. Crash avoidance technologies are evolving rapidly toward
increasing automation, involving a higher complexity of interoperability between user and vehicle functions than has previously been known. Understanding the detailed human factors capabilities and limitations of these users and the impacts of the timing, sequence, and presentation of information presented to the users will be important for shaping the safety policies.

REFERENCES

