Yury Park
CS7633 – Human-Robot Interaction

Written Case Study #3

I. *Relevant Literature and Key Insights*

A challenging paradox implicit in our research question is the fact that we are attempting to find ways to temper one of the key features that make autonomous driving attractive to many people in the first place: the convenience and freedom afforded to drivers to engage in other tasks while the vehicle is in motion. Various studies, including a recent one by Howard and Dai, have found that convenience and multitasking have been cited by a wide range of demographic (not correlated to any specific group) as major benefits of the autonomous driving technology.[1]

Surely, it isn't difficult to imagine one day when autonomous vehicles have been thoroughly vetted for superior safety and efficiency, and the current thorny debates pertaining to issues of liability (e.g. who takes the blame in the event of an accident involving an autonomous vehicle?) and ethics[2] (e.g. should the vehicular AI take on a utilitarian or deontological approach regarding the concept of minimization of harm?) have largely become a relic of the past. Furthermore, there already exist forward-thinking proposals that look beyond the horizon of self-driving vehicles (such as one by Gerla et. al. that seeks to combine autonomous vehicles along with intelligent road grid networks equipped with sensors / RFID tags to form what's called the Internet of Vehicles, which would use real-time collaborative Vehicular Cloud technology to help ease congestion, save records of accidents, allow for orderly coordination of traffic in the event of natural disasters, and so on[3]).

Currently, however, we are in the nascent stages of a fast-approaching transition period in which conventional and self-driving vehicles will most likely share the road, with the ratio between the former and the latter gradually shifting over time. The transition is happening in other ways as well. For example, the choice between manual and self-driving isn't binary, but rather comes in gradations; the degree of a vehicle's autonomy has been classified by the National Highway Transportation Safety Administration (NHTSA) into Levels 0 (fully manual) thru 4 (fully automated).[4] As of today, Level 2 (limited automation with steering, braking and lane guidance) vehicles are available in the consumer market, with Level 3 (restricted self-driving) vehicles undergoing testing.[5] For the foreseeable future, therefore, the advent of self-driving vehicles will be accompanied by the requirement that human drivers take on the role of supervisors, monitoring the road conditions and being ready to assume manual control at any time.

And it is this requirement that is exceedingly challenging to reconcile with one of the key selling points of self-driving technology (as we mentioned at the beginning). Numerous studies have consistently found that complacency and/or reduced focus and situation awareness are associated with autonomous driving.[6] For instance, an experiment by Strand, et. al. which employed an automated driving simulator (with various automation levels) discovered that higher level of automation negatively affected driving performance in situations when automation suddenly failed and the human driver needed to resume control on the fly.[7] Sometimes, the problem surfaces in the reverse fashion: accidents can happen not only because drivers of autonomous cars fail to intervene in time when they should, but also because they mistakenly intervene when there is no danger of the autonomous system's causing an accident.[8] In fact, it has been called into question whether a semi-autonomous approach requiring the continuous

---

[1] Howard & Dai. "Public Perceptions of Self-Driving Cars: The Case of Berkeley, California" (2013).
[2] *See e.g.* Hevelke & Nida-Rumelin. "Responsibility for Crashes of Autonomous Vehicles: an Ethical Analysis" (2014).
[3] Gerla, et. al. "Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds" (2014).
[4] NHTSA. "Preliminary Statement of Policy Concerning Automated Vehicles" (2013).
[5] Litman. "Autonomous Vehicle Implementation Predictions" (2015).
[6] *See e.g.* Goodall. "Machine Ethics and Automated Vehicles" (2013).
[7] Strand, et. al. "Semi-automated versus Highly Automated Driving in Critical Situations" (2014).
[8] Douma & Palodichuk. "Criminal Liability Issues Created by Autonomous Vehicles" (2012).

engagement of human drivers is even sustainable[9], since the more inured to the technology drivers become, the less likely they are to stay engaged to the road.

We sympathize with and share these concerns about the long-term sustainability of the quest to keep drivers engaged as vehicles progressively become more autonomous and safer. Nonetheless, the possibility that our task may not be viable in the long run in no way diminishes its importance during this critical transition period; after all, how we collectively handle the transition, in terms of minimizing confusion, complacency and ultimately accidents, may help determine whether there even is a long-term future for automated vehicles. Particularly, in a state such as ours where the driver of an autonomous vehicle is ultimately liable should an accident occur, it is exigent to design self-driving vehicles with a robust and intuitive driver-side interface that will consistently keep drivers alert and situationally aware of the road conditions, without unduly scaring them with numerous false alarms and warning signals.

To this end, we take inspiration from related studies such as one performed by Merat et. al.[10], in which drivers' attention span was measured in automated vehicle simulators under two situations: 1) automation / manual mode were toggled on and off at regular, system-based intervals, and 2) transition from automated to manual mode was based on the length of time that drivers were looking away from the road ahead. It was found that drivers' ability to regain control of the vehicle was better and more stable in the first situation – i.e. when drivers could expect the automation to be switched off. Our proposed hypothesis and experiment, to be described in the following sections, goes further to address the question of how best to keep drivers in the loop without resorting to the impractical method of turning off automation at fixed intervals – a question that, to our knowledge, has not been resolved by other studies.

II. Technical Approach

The research question may be decomposed as follows: 1) implementing functionalities and layout, and 2) measuring levels of driver engagement / performance upon takeover. We take functionality and layout to mean the design of features specific to self-driving vehicles, such as driver's eye movement trackers and notification systems, that are intended to assess or improve driver engagement. Driver engagement refers to the alertness level of the driver during automation, as may be indicated by the driver's visual focus on the road ahead, the ability to resume manual control safely and efficiently, and ability to prevent accidents in sudden and/or unexpected situations, among others.

As a brief overview, we list the two broad key questions to which we will be seeking answers in the course of our research. First, how do different types of automation affect levels of driver engagement? Second, what features should be implemented to maintain driver engagement over time and stave off complacency as well as false alarms, found to be correlated with humans' engagement in monitoring tasks?[11]

At a high level, we will begin by performing an analysis of existing studies and their data to better inform our baseline understanding of the above issues. Afterwards, we will form a tentative hypothesis (to be discussed further Secs. III and IV) and employ a series of experiments based on an automated vehicle simulator similar to the University of Leeds Driving Simulator (UoLDS) used in related studies,[12] to measure the impact that various changes in vehicle features and design have on drivers' alertness and readiness to take over manual control.

III. Analysis of Existing Data

[9] Stone, et. al. "One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel" (2016).
[10] Merat, et. al. "Transition to Manual: Driver Behaviour when Resuming Control from a Highly Automated Vehicle" (2014).
[11] *See* Parasuraman and Manzey. "Complacency and Bias in Human Use of Automation: an Attentional Integration" (2010).
[12] *See e.g.* Jamson, et. al. "Behavioural Changes in Drivers Experiencing Highly-Automated Vehicle Control in Varying Traffic Conditions" (2013).

As mentioned previously, existing studies have largely shown that humans in general perform poorly as monitors of automated technology.[13] There have also been studies suggesting that the *types* of automation (longitudinal vs. lateral control) provided to drivers may result in different levels of driver engagement.[14] As such, we will perform a more detailed analysis of existing literature and experimental data to form a preliminary hypothesis on the issue – namely, the impact that different levels (we will mainly consider Level 2 and Level 3[15]) and types of automation (longitudinal control, i.e. acceleration / deceleration, vs. lateral control, i.e. steering / lane keeping system), has on the driver's engagement and alertness, utilizing such metrics as i) reaction time to resume manual control, ii) longitudinal and lateral driving performance (measured by mean and minimum values of speed for the former, Standard Deviation of Lane Position (SDLP) and High Frequency Control of steering for the latter[16]) upon resuming manual control, and iii) eye tracking measure of Percent Road Center (PRC) which has been demonstrated to be a reliable indicator of visual distraction.[17] Our expectation, based on current research as mentioned above, is that higher levels of automation, combined with automated longitudinal and lateral control, will be correlated with decreased overall engagement and increased distraction on the part of the driver.

That brings us to the more salient question of feature implementation and design to help offset driver distraction and maintain situational awareness to guard against system failures or unexpected events. Existing research provides us with a number of insights on this issue. For one, while highly automated driving in general results in reduced driver reaction time and higher rate of collisions (or near-collisions) in critical scenarios when compared with manual driving, these results are affected in part by whether timely warning was provided; specifically, unexpected automation failure with little time for the human to respond causes almost all drivers to crash, whereas timely warning allows most drivers to avoid collisions.[18] Moreover, empirical data has suggested that automated driving can actually result in an improved situational awareness over manual driving, provided that drivers are instructed (and/or self-motivated) to detect objects in the surrounding environment as opposed to engaging in other, non-driving tasks.[19]

From these and other related work, including the aforementioned study by Merat et. al. (see Sec. I), we would like to form the following tentative hypothesis: we expect driver engagement (e.g. attention to the road ahead, reaction time / performance when resuming manual control, etc.) to improve in automated systems equipped with features that i) remind drivers to monitor the environment, ii) warn drivers in timely fashion of impending conditions or situations that require human intervention, iii) are robust in checking for and guarding against unexpected automation failures through multiple layers of built-in redundancy, and iv) employ vehicle-to-vehicle communication systems[20] to reduce the number of false alarms regarding collision avoidance. In seeking to validate this hypothesis, we plan to focus mainly on i) and ii) as they are most relevant to a HRI study, whereas iii) and iv) pertain more to the engineering / algorithmic domain, with the latter being only effective when other vehicles adopt the same or similar strategy, which we cannot assume for purposes of our research. The details of our validation are described in the following section.

IV. Evaluation and/or Validation

The validation of our hypothesis i) and ii) mentioned in the foregoing section will be performed through a series of experiments that employ a sophisticated driving simulator similar to the University of Leeds Driving Simulator (UoLDS), with eligible participants selected from a pool of experienced, capable drivers

[13] Strand. "Semi-automated versus Highly Automated Driving" (see footnote 7). *See also* Sheridan. "Humans and Automation: System Design and Research Issues" (2002).
[14] Carsten, et. al. "Control task substitution in semi-automated driving: Does it matter what aspects are automated?" (2012).
[15] NHTSA. "Preliminary Statement of Policy" (See footnote 4).
[16] *See* Merat. "Transition to Manual" (footnote 10) for an example of these metrics in use.
[17] Victor, et. al. "Sensitivity of Eye-Movement Measures to In-Vehicle Task Difficulty" (2005).
[18] Gold, et. al. "Influence of Automated Brake Application on Take-over Situations" (2014).
[19] Winter, et. al. "Effects of Adaptive Cruise Control and Highly Automated Driving" (2014).
[20] *See* U.S. Dept of Transportation. "Precursor Systems Analyses of Automated Highway Systems" (1994).

(e.g. minimum of 5 years of driving experience, with 20,000+ miles driven per year on average). The broad aim of the experiment is twofold:

1. What type of system reminder (regular vs. irregular), best improves driver engagement and alertness?
2. Does "priming" the drivers with reminders as per above result in improved performance in the event of sudden, unexpected events requiring immediate takeover?

The simulator will be equipped with real-time position, head and eye tracking sensors in order to assess the driver's level of attention to the road ahead. The driver-side seat will be equipped with a vibration mechanism to be used only in urgent situations to alert drivers to take over control. Above the speedometer panel, there will be a LCD screen showing whether the vehicle is in manual or automated mode, and capable of displaying / speaking other short messages to the driver for purposes of reminders and warnings. The steering wheel will have a sensor to verify whether the driver's hands are touching it. The longitudinal and lateral controllers will be, respectively, an Adaptive Cruise Control (ACC)[21] and a Lane Keeping System (LKS).[22]

Prior to the experiment, drivers will have a 20-minute training period to get familiarized with the simulator, and will learn of the purpose of the LCD, sensor and seat vibration features among others. Drivers will then be told of the simulated driving environment, which will be a 30-minute drive on a 3-lane highway, and instructed to manually drive the car for the first 3 minutes, after which automation will take over. They will be further advised in advance that they should monitor the road conditions during automation as they may need to take over control of manual driving on the fly due to various unexpected road conditions, and that two minutes after they have taken over control in such conditions, the automation mode will be turned back on. Details that will not be divulged to the drivers ahead of time include the fact that different drivers will receive different types of reminders (or no reminders), and that they will be required to take manual control for a short while both at the 15-minute mark (merging onto another lane due to construction work sign) and 25-minute mark (decelerating or avoiding rapidly due to metal shards / other small but dangerous debris on the road ahead that the automated system cannot identify for certain and requests driver intervention). We believe that letting the automation take over for a duration of 10 – 15 minutes each without incident will make it more likely for complacency or fatigue to set in, allowing us to better measure the impact of unusual or sudden events on the driver performance.

This will be a between-participants design study to ensure the separation of treatment and baseline groups. The dependent variables will consist of the drivers' longitudinal and lateral driving performance as well as head and eye tracking measurements for the first 30 seconds after each of the two times they are required to take over manual control. The independent variable will be the type of system reminders (regular / irregular) provided, with some drivers not receiving any reminder. Regular system reminders will be provided every five minutes irrespective of drivers' level of alertness, and will consist of visual displays on the LCD above the speedometer along with spoken language, and will be a statement similar to "Please keep your eyes on the road ahead and hands on the steering wheel at all times", or "unexpected situations can arise at any time. Please continue observing the road". Irregular system reminders, on the other hand, will be provided only if the tracking system detects that the drivers have been looking away from the front and/or at least one hand has not been touching the steering wheel for over ten seconds. Some of the subjects will receive the regular reminders, others the irregular reminders, and the baseline group will receive no reminders at all.

In addition, visual / spoken system notifications will be provided to all drivers at the 15- and 25-minute marks. The former notification will concern the construction sign mentioned above, and be in the form of "There appears to be a custom sign ahead. Please resume manual control if needed." The driver will then be expected to manually merge onto a different lane. The latter notification will concern the dangerous debris on the road and will state "Warning: possible obstacle detected ahead," along with a vibrating seat to further alert the user to the urgency of the situation, and the driver will be expected to manually avoid

[21] As modelled in the study by Ioannou, et. al. "Intelligent Cruise Control: Theory and Experiment" (1993).
[22] As modelled in the study by Casanova & Symonds. "A Mathematical Model for Driver Steering Control" (2000).

the debris. After the experiment, the subjects will be presented with a survey to gather their subjective states of mind during the simulation, e.g. their level of alertness / boredom, whether they found certain types of reminders helpful or annoying, and how they reacted when they were expected to take over control.

Statistical analysis will be performed to compare driver performance and engagement for the drives with regular, irregular, and no reminders. A 1-way ANOVA with the aforementioned factors will be performed for each of the following specific metrics: 1) the length of time drivers' gaze diverted away from the road ahead during the drive, as measured by the eye and head tracking sensors; 2) the length of time drivers' hands were taken off the steering wheel; 3) longitudinal and lateral driving performance for 30 seconds after taking manual control at the 15- and 25-minute marks (as measured by the mean values of speed and Standard Deviation of Lane Position (SDLP[23]), respectively); 4) the eye tracking measure of PRC[24] for the 30 seconds after taking manual control in both situations. Afterwards, we may perform a post-hoc analysis to look for the existence of any interactions between the above metrics. Our present expectation is that the presence of reminders (whether regular or irregular) will result in improved driver engagement and performance over the absence of reminders, and that regular reminders will be more effective than the latter. We will revise our hypothesis as necessary depending on the result of the experiment.

Finally, we may perform a second set of experiments to further validate the findings of the study by Gold et. al.[25] and to identify any interaction between reminders and timely notifications of unexpected events. This set of experiments will be identical to the former, except that some drivers will not be provided with any system notification at either the 15- or 25-minute mark. We expect to find that the presence of reminders will help stunt the degree of negative impact that the lack of system notifications will have on driver performance in unexpected or dangerous situations. Future area of research, time permitting, may include an examination of 1) the possible desensitizing effects of drivers' repeated exposure to the same type of reminders and whether this results in a degradation of performance over time, and if so, 2) ways to mitigate the issue through injecting variations into the wording of the reminders depending on the sensors' perception of the traffic conditions, and/or requesting driver feedback in the form of oral or visual confirmation.

[23] See, e.g. McLean & Hoffman. "The Effects of Restricted Preview on Driver Steering Control and Performance" (1973).
[24] Victor. "Sensitivity of Eye-Movement Measures" (See footnote 17).
[25] Gold, et. al. "Influence of Automated Brake Application on Take-over Situations" (See footnote 18).