IMPROVING DRIVER ENGAGEMENT DURING L2 AUTOMATION: A PILOT STUDY

Anuj K. Pradhan
University of Massachusetts Amherst, Amherst, Massachusetts, USA
anujkpradhan@umass.edu

Jacob Crossman
Soar Technology, Ann Arbor, Michigan, USA
jcrossman@soartech.com

Adam Sypniewski
Deepgram, San Francisco, CA, USA
adam@deepgram.com

Summary: Advanced technologies such as adaptive cruise control and lane keeping are key components of SAE Level 2 vehicle automation. As such automation becomes widespread, drivers may be less engaged in driving because they assume that vehicles can safely mitigate risks. However, L2 automation cannot handle the full spectrum of driving situations and will require manual control in many situations. Drivers unprepared to take control may make suboptimal, delayed, or dangerous decisions during and after reengaging with the driving task. This highlights the need for efficient ways to help drivers re-engage with driving. This paper describes an evaluation of a conceptual driver engagement system that combined driver data with contextual data to communicate appropriate information during L2 operations. The system was compared to a traditional, staged-alert system that only monitored driver gaze with no contextual information. Results indicate higher situation awareness, higher levels of trust and satisfaction, no increase in workload, with evidence of improve off-road glance behaviors when driving with the conceptual system. These findings can help inform further development and testing of driver engagement approaches using driver monitoring.

BACKGROUND

Auto manufacturers have made considerable progress in developing advanced driver assistance systems and precursors to high levels of vehicle automation, and have released technologies that in combination can be categorized as Level 2 (L2) vehicle automation as defined by the Society of Automotive Engineers (SAE) (SAE, 2018). L2 automation may provide sustained automatic lateral and longitudinal control of a vehicle, but the driver is still required to maintain attention to the driving task and should be capable of taking over control of the vehicle immediately if L2 systems become incapable of maintaining control. Current examples of such vehicle technologies include Adaptive Cruise Control (ACC) and Lane Keeping Assistance (LKA). These systems can often work in concert taking away from the driver the need to manage vehicle longitudinal and lateral control. These technologies are now readily available in late model vehicles.

Research on automation has shown that when humans who are out-of-the-loop need to take back manual control (Wickens & Kessel, 1979) there are associated resumption costs. Similarly, there is contemporary recognition in the automotive human factors research community of the challenges involved in the transition between automation and the driver (Jamson et al., 2013; Gold et al., 2013; Merat et al., 2014). For developers of automated vehicles, raising automation levels to those where human intervention is not required may be the ultimate goal, but while there are situations that automation cannot handle the driver remains a crucial part of driving, such that "driving safety … depends on the combined performance of the human and automation" (Merat & Lee, 2012). The obvious concerns are safety-related, with increased automation having the potential to result in delaying of driver response (Neubacher et al., 2012) in circumstances where intervention in the form of manual takeover of control is expected.

OBJECTIVE

One approach towards addressing safety related concerns with automation is by ensuring that drivers are monitored by in-vehicle sensors for driver attention state and somehow kept engaged in the driving task regardless of automated control. Driver engagement may reduce the likelihood that the system will reach its limits without early human intervention, and, if these limits are reached, can ensure that the driver's awareness is sufficient for regaining control (Gold et al., 2013; Merat et al., 2014). However, a balance needs to be maintained between convenience of automation versus driver engagement. The challenge is not simply to keep the driver engaged, but rather to engage the driver while still retaining the conveniences of automation.

Thus, the objective of this study was to evaluate, as a pilot step, a conceptual system designed to leverage driver state monitoring as an approach towards keeping drivers engaged during L2 automation. The conceptual system (System A) was designed as a driver assistance sub-system using both driver and environmental characteristics. This system was experimentally evaluated by comparison with a more traditional approach (System B) towards driver engagement using gaze metrics without contextual data. The evaluation was undertaken in a driving simulation platform that permitted the prototyping of the systems and the simulation of L2 automation.

METHODS

System A description. The goal was to enable the advantages of L2 driving, convenience and safety, while maintaining the driver engagement required for safe driver takeover from automation. The system was therefore designed to adaptively engage the operator in supervisory control of the vehicle based on a model of the user and the environment to determine when and how to interact. The user model was informed by driver gaze data, and the environment model was informed by the driving scenarios and context. The system was designed to interact with the driver using multiple modes including visual, audio, voice, and haptic feedback. Specifically, for System A, the simulator was equipped with four different modalities to interact with the driver: haptic actuators under driver seat, a multi-color LED light bar above the dashboard, a head-up display (HUD) (Figure 1), and audio (warning chimes and voice synthesis).

During the drive, at various scenarios (pre-programmed events), contextual information was displayed on the HUD describing the upcoming scenario and indicating how the vehicle automation would behave during the scenario. If the driver was not paying attention (did not spend at least 3 consecutive seconds looking at the road in a 30 second interval), or if take over was required, then a warning chime, a spoken description of the scenario, and haptic feedback were activated. If the driver continued to ignore the road, the light bar would flash red, the haptic feedback was activated, and a voice directed the driver to take control immediately. Situational awareness and re-engagement alerts were also issued during periods of prolonged distraction between scenarios. These non-scenario alerts used the HUD, haptic feedback, and voice to interact with the driver. Example non-scenario alerts include, “Highway entrance ramp ahead; please scan for traffic,” “Possible speed limit change ahead; please watch for speed limit signs,” and “You have been distracted for a long time; please scan for traffic.”

Figure 1: HUD & LED light bar

System B description. The comparison system, System B, was based on concepts described in Blanco et al.. (2015). Here, the drivers’ gaze was monitored in real time. Based on gaze, if the driver was inattentive for more than 7 seconds, the light bar flashed yellow. If the driver continued to remain inattentive for an additional 5 seconds, the light bar flashed red and haptic feedback was activated. If the driver still remained inattentive for an additional 5 seconds, the light bar flashed red, the haptic feedback was activated, and a warning tone was played back – after which the driver was required to take back control of the vehicle. There was no information provided on the HUD, and no contextual information provided to the driver about the scenarios.

Driving Simulator. The evaluation of the systems was conducted on a high-fidelity, fixed based driving simulator (Realtime Technologies Inc.) The simulator comprised a full-cab Nissan Versa equipped with a SmartEye four-camera eye-tracking system that provided real time head pose, eye blink, and gaze data. The simulator immersed the participant in the virtual driving environment via three screens in the forward direction, plus one in the rear, and two virtual side mirror displays. The simulation platform was further programmed to simulate various levels of automation. In this study, the simulator was programmed for L2 automation. The vehicle started under complete driver control, then the driver transferred longitudinal and lateral control to the automated system, and control was transferred back to the driver at pre-determined situations.

Experimental Design, Participants, & Protocol. The experiment was designed as a within-subject study. The two systems were implemented in the simulation using a Wizard-of-Oz (WoZ) approach. The experimenter manipulated alert time and modalities based on real-time observations of the participants using a custom designed control panel integrated with the simulator systems. This panel helped execute the WoZ in a controlled manner. Upcoming scenarios were displayed in real-time on the panel, along with driver gaze & attention information. The panel provided response options to display appropriate alert combinations to the driver. Seven participants were recruited and were presented with two fifteen-minute simulated highway drives. Each drive contained ten scenarios. Driver awareness was monitored using gaze as an indicator of attention. If drivers remained inattentive for too long, alerts (described above) were provided via WoZ to re-engage attention. Either systems A or B were presented in each drive, with presentation order counterbalanced across participants. In both drives, participants were asked to conduct secondary tasks (simple iPad tasks such as watching short videos or looking up information). Participants were instructed to engage automation after starting the drive, and then to disengage when they thought it appropriate, and re-engage automation after each disengagement once they felt comfortable doing so.

Outcome measures: There were two categories of outcome measures, surveys and eye gaze. Three surveys were presented to the drivers after each drive measuring satisfaction, trust, and workload. The satisfaction survey asked about ease-of-use and usefulness of the system. The trust survey probed confidence in system performance. The workload survey was a standard NASA-TLX survey (Hart & Staveland, 1988). A single situation awareness survey was also administered after both drives were completed. This survey asked how aware drivers were of a scenario before it occurred, for each drive. Eye-movement measures included gaze location and duration recorded at 60 Hz during the drives via the integrated eye-tracking system.

RESULTS

L2 Usage. Analyses of participants' use of L2 during the drives showed that they spent more time in L2 with System A (81%) than with B (67%), although the difference was not significant.

Satisfaction. This survey probed for overall satisfaction with the system's performance, including questions about satisfaction with the number and type of alerts, having sufficient time and information to make decisions, and ease of use. General trends in subject satisfaction were measured by pooling survey responses. Treating the response data as interval (on a scale of 1–7), a one-sided paired t-test indicates that System A ($M = 5.99$, $SD = 1.03$) scored significantly better than System B ($M = 5.07$, $S = 1.68$), $t(69) = 4.57$, $p < 0.001$, in the satisfaction survey.

Trust. This survey probed subjects about overall trust in the system, including questions about reliability and dependability, providing alerts when needed, false alert rates, comfort with taking over control, and agreeing that takeover was necessary when prompted to do so. Survey responses were pooled to show participants reporting significantly higher trust scores with System A ($M = 5.79$, $SD = 1.11$) as compared to B ($M = 5.09$, $S = 1.60$), $t(55) = 3.72$, $p < 0.001$.

Workload. Workload was measured using a NASA-TLX survey, with both scale and weight sections. Overall workload scores, calculated as the mean of the sum of the weighted workload scores, were the same for Systems A (M = 44.3) and B (M = 42.1). A two-way repeated measures ANOVA showed no interaction between the workload traits and the alert system, $F(5, 30) = 1.3472$, $p = 0.27$. Paired t-test comparisons with Holm-Bonferroni corrections between System A and System B for each workload trait also showed no significant differences.

Situation Awareness. The situation awareness survey prompted subjects about each of the ten scenarios. For each scenario, they were presented with two scales—one for each drive—and asked to rank their awareness of each scenario before it occurred. Participants reported significantly higher awareness when using System A (M = 6.54, SD = 1.18) as compared to B (M = 4.76, SD = 2.05), $t(66) = 6.78$, $p < 0.001$.

Gaze Duration. First, participants had similar off-road gaze durations for both systems, with slightly more off-road glances overall (54%) in System A, compared to B (48%). However, this was not statistically significant. Second, analyzing gaze only during automation, participants spent more total time looking off-road during L2 automation with System A (60%) compared to System B (52%), $\chi^2(1) = 67.38$, $p < 0.001$; however, when L2 was disengaged, participants spent less total time looking off-road in System A (27%) than in System B (40%), $\chi^2(1) = 61.68$, $p < 0.001$. Finally, on average, drivers spent significantly longer watching the roadway during the seconds before takeover with System A (77%) than B (58%), $\chi^2(1) = 33.58$, $p < 0.001$.

Gaze Dispersal. The participants' gaze dispersal during the drives can be visualized using contour plots of their gaze angles. Figure 2a and 2b show the density of gaze angles in the forward direction for Systems A and B respectively. Blue indicates regions with low gaze density, orange and white indicate those with high gaze density. The origin (0, 0) is straight ahead down the road, and the HUD is located approximately at (0,−5). The region at (40, 10) is the rear-view, at (−35,−15) the driver side, and at (60,−10) the passenger side mirror. The region around (0, −40) is below the steering wheel. In both drives, drivers looked at obvious drive-related gaze points: roadway, mirrors, HUD, and the secondary task. Figure 2c visualizes the difference between the two heat maps to show differences in gaze patterns, with the red area indicating areas more commonly looked at during System A drives, and blue showing areas more common during B. The only significant difference in gaze patterns seem to be towards the HUD in System A, which is active during System A drives.

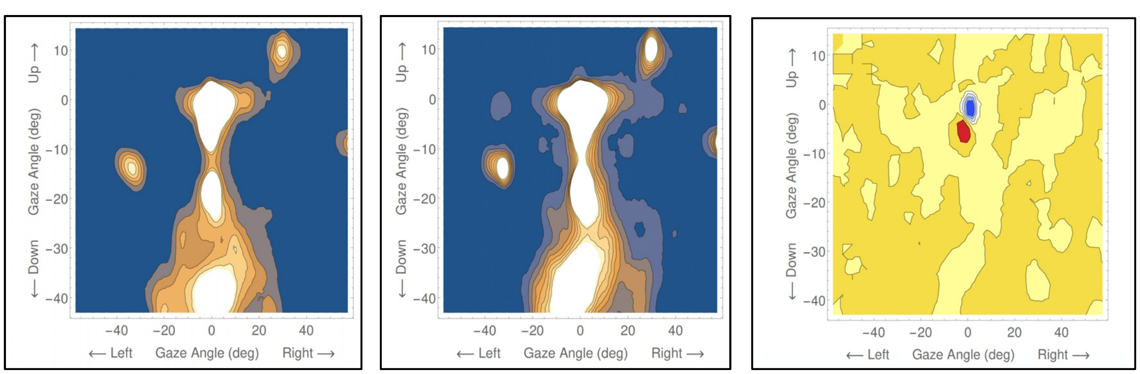

Figure 2: (a) System A; (b) System B; (c) System A & B difference

CONCLUSION

The results indicate that participants tended to spend more time (14 percentage points) in L2 mode when using System A, than when using System B. One interpretation is that this may have been driven by the drivers' satisfaction and trust in the automation since survey results indicate higher satisfaction with the System A alert protocols than with System B, and higher trust in the L2 system when driven with System A than with System B. Given that the design intent of System A was to foster driver engagement while preserving convenience and minimizing 'nagging' alerts, these results seem to suggest that such a goal may indeed have been achieved. In addition, there were no significant differences in driver workload between the two systems, indicating that added modalities of alerting in System A as well as the continuous information provided, say by the HUD, did not necessarily add to the overall driving workload.

The situation awareness survey responses suggest that participants remained more aware of the driving situation when using System A as compared to the more traditional alert system. The glance data further reinforce this for the takeover situations. Since System A was specifically designed to provide situation awareness through multiple modes, including spoken contextual information that offered more detailed descriptions of each situation, this finding seems unsurprising, even for the small sample size. Despite participants tending to have more off-road glances during automation when using System A, the spoken descriptive of scenarios may have contributed to increased situation awareness despite longer off-road glances in System A, whereas driving with System B did not provide that level of information about the environment.

The eye tracking analyses show that drivers' overall visual attention and gaze patterns were very similar when using the two systems. Overall, there were no differences in the duration of off-road (non-driving related) glances. Additionally, drivers' scan patterns were nearly identical between the two systems (Figure 2). The one important difference was that drivers using System A spent less time looking at the road overall and when automation was engaged. This particular finding can potentially be addressed by two explanations: First, the System A concept used a different alert protocol than System B, wherein, the former warned the driver to look at the road if they did not spend at least 3 consecutive seconds looking at the road in a 30 second interval, while in the latter, the driver was warned after 7 consecutive seconds of inattention. This difference may have resulted in the driver being able to look away from the road longer before being alerted in System A. Second, the surveys indicate the drivers trusted L2 more and believed they were more aware when using System A than not, which may have contributed to the driver being more comfortable with more off-road glances during highway driving. Despite this above finding, which indicates a somewhat diminished attention to the forward roadway with System A, the results also do show that drivers using System A seemed to spend more time looking at the road before the actual takeover situations, a more critical time to be aware of the environment.

This study was conducted as a pilot simulator evaluation of driver responses to a prototype of a conceptual driver engagement approach. An important limitation, however, is the small sample size in this study, seven participants recruited for a within subject experiment. The sample size was limited due to time and resource constraints. Although this pilot was ultimately informative in providing information for further development and had some compelling findings, a larger sample size may have had provided more power lending more confidence in the results.

Nonetheless, the findings from this pilot evaluation valuable. They have indeed informed further development efforts, including ongoing work on the building of a functional prototype of this system for integration on a live vehicle, and ongoing evaluation on a test track. These results can thus help inform development of advanced driver assistance systems (ADAS) that can monitor the driver and environment, with intelligent and adaptive driver engagement approaches, helping retain promised safety benefits of ADAS without additional human factors related risks.

ACKNOWLEDGMENTS

This work was supported by funding provided from the US Department of Transportation's Federal Highway Administration through the US DOT's Small Business Innovation Research program.

REFERENCES

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., ... & Morgan, J. F. (2015). *Human factors evaluation of level 2 and level 3 automated driving concepts* (No. DOT HS 812 182).

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013, September). "Take over!" How long does it take to get the driver back into the loop?. In*Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(Vol. 57, No. 1, pp. 1938-1942).

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183).

Jamson, A. H., Merat, N., Carsten, O. M., & Lai, F. C. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies*, *30*, 116-125.

Merat, N., & Lee, J. D. (2012). Preface to the Special Section on Human Factors and Automation in Vehicles: Designing Highly Automated Vehicles With the Driver in Mind. Human Factors, 54(5), 681–686.

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation research part F: traffic psychology and behaviour*, *27*, 274-282.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(5), 734-746.

SAE (2018). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles – J3016_201806.

Wickens, C. D., & Kessel, C. (1979). The effects of participatory mode and task workload on the detection of dynamic system failures. *Systems, Man and Cybernetics, IEEE Transactions on*, *9*(1), 24-34.