2022 NDIA MICHIGAN CHAPTER
GROUND VEHICLE SYSTEMS ENGINEERING
AND TECHNOLOGY SYMPOSIUM
MODELING SIMULATION AND SOFTWARE (MS2) TECHNICAL SESSION
AUGUST 16-18, 2022 - NOVI, MICHIGAN

DEVELOPING A MODEL OF DRIVER PERFORMANCE, SITUATION AWARENESS, AND COGNITIVE LOAD CONSIDERING DIFFERENT LEVELS OF PARTIAL VEHICLE AUTONOMY

Jessie E. Cossitt, PhD[1,2], Viraj R. Patel,[1,2], Daniel W. Carruth, PhD[2], Victor J. Paul[3], Cindy L. Bethel, PhD[1,2]

[1]Computer Science and Engineering, Mississippi State University, Starkville, MS
[2]Center for Advanced Vehicular Systems, Mississippi State University, Starkville, MS
[3]U.S. Army CCDC-Ground Vehicle Systems Center, Warren, MI

ABSTRACT

To optimize the use of partially autonomous vehicles, it is necessary to develop an understanding of the interactions between these vehicles and their operators. This research investigates the relationship between level of partial autonomy and operator abilities using a web-based virtual reality study. In this study participants took part in a virtual drive where they were required to perform all or part of the driving task in one of five possible autonomy conditions while responding to sudden emergency road events. Participants also took part in a simultaneous communications console task to include an element of multitasking. Situation awareness was measured using real-time probes based on the Situation Awareness Global Assessment Technique (SAGAT) as well as the Situation Awareness Rating Technique (SART). Cognitive Load was measured using the NASA Task Load Index (NASA-TLX) and an adapted version of the SOS Scale. Other measured factors included multiple indicators of driving performance and secondary task performance. Results indicate a relationship between performance and autonomy level.

Citation: J. E. Cossitt, V. R. Patel, D. W. Carruth, V. J. Paul, C. L. Bethel, ”Developing a Model of Driver Performance, Situation Awareness, and Cognitive Load Considering Different Levels of Partial Vehicle Autonomy,” In *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium* (GVSETS), NDIA, Novi, MI, Aug. 16-18, 2022.

1. INTRODUCTION

The motivation behind this work is to create the optimal utilization of autonomous vehicle capabilities and to ultimately create a system of dynamic allocation of tasks to reduce the necessary crew size on military missions [1]. Because the current state of the art in autonomous vehicles is partial autonomy that requires the operator to be alert and ready to respond to sudden events, operator abilities have to be taken into account.

Based on previous research it would be reasonable to expect increasing levels of partial vehicle autonomy to have a negative impact on situation awareness and cognitive load [2] [3]. However, there are many factors influencing the driving task, especially in a military context. These factors include the use of autonomy to allow for multitasking and the balance between cognitive underload and overload of the operator [4].

Developing a function of operator ability at multiple distinct levels of partial vehicle autonomy while taking into account the use of autonomy to allow the operator to also perform other mission tasks should lead to the development of better autonomous systems designed with the operator in mind. These systems would be able to predict operator ability and behavior and take measures to increase engagement in scenarios where abilities are predicted to decline. Additionally, this knowledge is important to create a better understanding of what should be required of the autonomous system in situations where operator takeover is necessary.

2. RELATED WORK

A study by Anderson et al. [5] investigated the possibility of reducing crew sizes using autonomous vehicle abilities through the use of static allocation methods. This study specifically reduced a three-person crew to a two-person crew, and the methods used in this work relied on the sharing of information between the two crew members.

A closed hatch vehicle study looking at military tasks was conducted by Metcalfe et al. [6]. This study focused on non-driving tasks of passengers in a closed hatch environment and investigated the soldiers situation awareness and cognitive load. The task that was used in the study required participants to detect potential threats in the environment. Participants' performance in this study indicated that in such tasks, crew members could become quickly overwhelmed and unable to perform when receiving too much information likely due to limits in short term memory.

Another military task study by Hollands et al. [7] looked at different types of tasks and how they affected situation awareness and cognitive load. The tasks that were compared were those that were based on auditory commands and those that were based on visual commands. The results of the study showed that auditory commands had less negative effect on situation awareness and cognitive load than visual commands. The study additionally looked at the effect of task rate and found that a more rapid rate of commands has more negative effect on situation awareness and cognitive load than a slower rate of commands.

Outside the military domain research has been done looking at normal driving conditions and handover requests issued by the partially autonomous vehicle. A study by Miller et al. [8] looked at the effects of different autonomous abilities on driving behavior. This study found that some kinds of autonomous abilities have more pronounced effects on performance than others.

3. EXPERIMENT DESIGN

For this experiment 180 participants were recruited for online participation using Amazon

Mechanical Turk. Each participant was randomly assigned to one of five possible autonomy conditions. The data of 30 participants was unusable due to the participants' failure or inability to complete all parts of the study. After the removal of this data, there were 150 usable data points, 30 for each condition.

Participants took part in the study by operating a virtual vehicle using keyboard and mouse controls. They first completed a three-minute familiarization session before taking part in a ten-minute experiment drive. During the experiment drive, participants operated in one of five possible autonomy conditions which each increased the autonomous ability of the previous condition by adding a new ability. These conditions are listed in Table 1. The first condition was the control with no autonomy. The second added longitudinal control (adaptive cruise). The third added lateral control (lane keeping). The fourth added automated turning at intersections, and the fifth added obstacle detection and avoidance for stationary obstacles. None of the autonomy conditions could respond to sudden events, and participants were instructed to take control if sudden emergency events occurred.

Table 1: Sample table.

Autonomy Condition	Vehicle Abilities
1	No Autonomy
2	Longitudinal Control (Adaptive Cruise)
3	Condition 2 + Lateral Control (Lane Keeping)
4	Condition 3 + Automated Turning at Intersections
5	Condition 4 + Obstacle Detection and Avoidance

The virtual environment (see Figure 1) for the study showed the participant a first-person view in a closed-hatch military vehicle with a steering wheel and indirect view screen directly in front of them. To the right was a communications console where participants received visual prompts of yes or no questions that they were instructed to respond to as quickly as possible while also safely operating the vehicle. These questions occurred at a constant rate of two questions per minute.

Figure 1: The virtual vehicle environment.

The environment outside the vehicle was made up of abandoned city streets with barricades creating only one possible way to navigate through intersections. Participants were told their task was to provide surveillance of an evacuated city. They did occasionally encounter other military vehicles on the streets.

3.1. Road Events

In the experiment drive participants encountered three different sudden road events requiring response. These events were randomly selected from the following:

- •A pedestrian running into the road.
- •A dog running into the road.
- •A truck pulling out in front of the vehicle.
- •Rubbish falling off a truck in front of the vehicle.
- •A barrel rolling into the road.
- •A ladder falling off of a building into the road.

Figure 2 shows an example of one of the road events where a pedestrian ran into the road in front of the vehicle. To respond to a sudden road event, participants had to disengage the autonomous abilities (if active) and swerve or brake to avoid collision.

Figure 2: An example road event with a pedestrian running into the road.

4. EVALUATION

Participants were evaluated during the experiment drive based on their driving performance, secondary task performance, and response to real-time situation awareness probes. After the experiment drive participants also responded to questionnaires rating their cognitive load and situation awareness as well as collecting demographic information.

Driving performance was based on multiple factors that were recorded during participants' drives. These factors included the following:

- Number of sudden road events hit
- Number of other vehicles hit
- Number of barricades hit
- Number of stationary obstacles hit
- Number of forced autonomy control swaps

These factors were combined into a overall hazardous driving score that was evaluated. Additionally the number of sudden road events hit was evaluated on its own using non-parametric testing.

Secondary task performance was evaluated based on the percentage of accurate responses, the average latency of responses, and an overall score that took both accuracy and latency into account. For this score each question was worth thirty points if answered in the first five seconds. After that, the question lost five points every five seconds down to a minimum score of one point per correct answer.

Situation awareness was assessed in two ways. The first was with responses to real-time situation awareness prompts that were built into the communications console task. These situation awareness prompts were based on Situation Awareness Global Assessment Technique (SAGAT) [9] methodologies assessing three levels of situation awareness: perception, comprehension, and projection. Unlike the original SAGAT recommendations, the prompts were given real-time as supported by research by Jones and Endsley [10]. The prompts were additionally visual rather than auditory, but since they were created to be very similar to the secondary task itself, they were not assumed to add any difficulty of their own.

The second metric of situation awareness was the Situation Awareness Rating Technique (SART) [11] which was given as a post-drive assessment. The SART is a subjective metric of situation awareness where participants rate their own situation awareness based on ten different questions with responses on a scale of one to seven.

Cognitive load was also evaluated in two ways using both NASA-TLX [12] and a modified version of the SOS Scale which was originally used for evaluating learning technologies [13]. Both metrics were given after the experimental drive and had participants rate a number of categories. NASA-TLX included six questions, each of which were responded to with a 21-point scale as in the original documentation. The SOS Scale questions that were modified for this study had responses in the form of a number between zero and 100. These questions included the following:

- "I consider the task of operating the vehicle at this moment:"
- "I consider the task of responding to tasks at the moment to be:"

•"I consider working with the system at this moment to be:"

•"The tools provided make the task at this moment:"

5. RESULTS

The factors of hazardous driving score, secondary task score, real-time situation awareness score, SART score, NASA-TLX score, and SOS Scale score were analyzed as part of a multivariate analysis of variance (MANOVA) for global effects of autonomy level. The results of this analysis were found to be statistically significant (F(4,145) = 2.33, p= 0.0004, α = .05) with a medium effect size (η_p^2 = 0.09).

Univariate ANOVA's of the contributing factors showed significant results for only hazardous driving score (F(4,145) = 12.31, p= 1.203e-08, α = .05) with a large effect size (η_p^2 = 0.25). Post-hoc testing of this factor using Tukey HSD [14] showed significant differences in pairwise comparisons of conditions one and two, conditions one and three, conditions one and four, conditions five and two, and conditions five and three. Figure 3 shows the distribution of hazardous driving scores for each condition

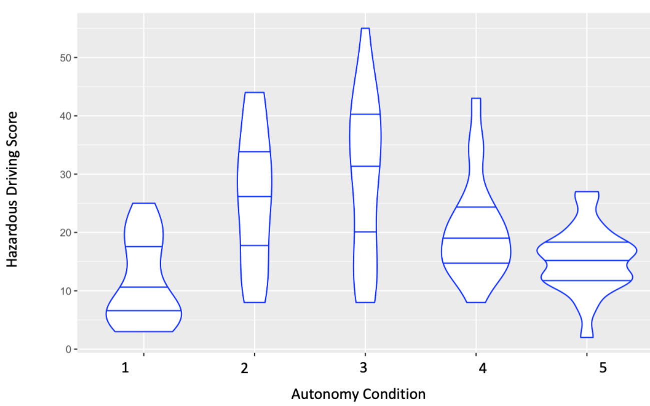

Figure 3: Violin plot of hazardous driving score for each autonomy condition.

Non-parametric testing of road event hits was also done in addition to overall score analysis. The Kruskal-Wallis rank sum test [15] was used to analyze the number of hits relative to autonomy condition. The test showed significant results with a p-value of 2.979e-07 and a chi-squared value of 35.94. The effect size was analyzed using eta squared based on the H-statistic and showed a large effect size (η^2[H]=0.22). Figure 4 shows the distribution of events hit in each condition.

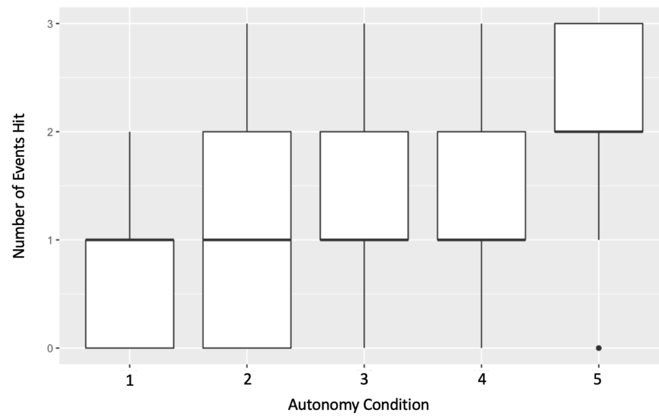

Figure 4: Box plot for events hit in each autonomy condition.

A Wilcoxon rank sum test [16] was used as a post-hoc analysis for pairwise comparisons of event hits for each autonomy condition and showed that all the significant pairwise comparisons were those involving condition five.

6. DISCUSSION

It is clear based on the results that driving ability was affected by autonomy level. An unexpected aspect of this result was that all significant differences in number of sudden events hit occurred involving condition five. This implies that the small amount of non-autonomous driving required in condition four to avoid obstacles was enough to increase engagement and avoid collisions with sudden events. The results of hazardous driving score analysis tells a different side of the story. This result shows worse driving performance in the middle three conditions. When further analyzing individual components of this score, it is clear that the increase in these conditions was caused by the factor of forced autonomy control swaps which could never occur in condition one and rarely could occur in condition five. This factor made up the smallest part of the score, but occurred frequently enough to skew the results on its own. This indicates that participants relied on the vehicle to disengage for them when encountering a situation outside the autonomous system's operating domain.

Other factors did not show significant results in this study. This could be potentially because of a lack of relationship but more likely due to a lack of intensity in the overall task. Subsequent studies will examine the factors with more intense tasks.

6.1. Future Work

Another experiment repeating what was done in this study with a steadily increasing rate of secondary task commands to find a breaking point has now been finalized and is in publication preparation. The combined results from these two studies will be used to create a detailed model of the interactions among the measured factors, task rate, and autonomy level. This model will then be used to create a system of dynamic task allocation for the use of autonomous vehicles to reduce crew size in military missions. A study will be done to test the developed system.

6.2. Conclusions

The main result found in this experiment was a finding that the operators of partially autonomous vehicles are incapable of responding to sudden emergency events when functioning in a purely supervisory driving roll. Other autonomy conditions where the operator shared in more of the driving task did not show a decrease in performance. Future work will further develop a model investigating changing task rates, and this model will be used to dynamically allocate mission tasks.

References

[1] J. E. Cossitt, C. R. Hudson, W. Carruth, Daniel, and C. L. Bethel, "Dynamic Task Allocation and Understanding of Situation Awareness Under Different Levels of Autonomy in Closed-Hatch Military Vehicles," *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium (GVSETS), NDIA, Novi, MI, Nov. 3-5, 2020*.

[2] M. Johns, S. Sibi, and W. Ju, "Effect of cognitive load in autonomous vehicles on driver performance during transfer of control," *AutomotiveUI 2014 - 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, in Cooperation with ACM SIGCHI - Adjunct Proceedings*, 2014.

[3] M. R. Endsley, "The limits of highly autonomous vehicles: an uncertain future: Commentary on Hancock (2019) Some pitfalls in the promises of automated and autonomous vehicles," *Ergonomics*, vol. 62, no. 4, pp. 496–499, 2019.

[4] M. A. Neerincx, "Cognitive task load design: model, methods and examples," *Handbook of cognitive task design*, p. 283–305, 2003.

[5] E. Anderson, J. Bryant, V. Paul, A. Reid, and J. Schrader, "USE OF CREW AUGMENTATION TECHNIQUES TO REDUCE COMBAT CREW SIZE," in *2018 NDIA Ground Vehicle Systems Engineering and Technology Symposium*, pp. 1–11, 2018.

[6] J. S. Metcalfe, D, K. A. Cosenzo, T. Johnson, B. Brumm, C. Manteuffel, A. W. Evans, and T. Tierney, "Human Dimension Challenges to the Maintenance of Local Area Awareness Using a 360 Degree Indirect Vision System," in *2010 NDIA Ground Vehicle Systems*

Engineering and Technology Symposium, 2010.

[7] J. G. Hollands, T. Spivak, and E. W. Kramkowski, “Cognitive Load and Situation Awareness for Soldiers: Effects of Message Presentation Rate and Sensory Modality,” *Human Factors*, vol. 61, no. 5, pp. 763–773, 2019.

[8] D. Miller, A. Sun, and W. Ju, “Situation awareness with different levels of automation,” in *EEE International Conference on Systems, Man and Cybernetics*, pp. 688–693, IEEE, 2014.

[9] M. Endsley, S. J. Selcon, T. D. Hardiman, and D. G. Croft, “A Comparative Analysis of Sagat and Sart for Evaluations of Situation Awareness,” *Paper presented at the 42nd Annual Meeting of the Human Factors and Ergonomics Society Chicago, IL — October, 1998*, 1998.

[10] D. G. Jones and M. R. Endsley, “Can Real-Time Probes Provide a Valid Measure of Situation Awareness,” *Proceedings of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference — October, 2000 CAN*, 2000.

[11] R. M. Taylor and S. J. Selcon, “Understanding situational awareness,” *Proceedings of the Ergonomics Society’s 1990 Annual Conference*, pp. 105–111, 1990.

[12] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” vol. 52, p. 139–183, 1988.

[13] J. Swaak and T. De Jong, “System control in using online support for simulation-based discovery learning,” *Learning Environments Research 4*, pp. 217–241, 2001.

[14] W. Haynes, “Tukey’s Test,” *Encyclopedia of Systems Biology. Springer, New York, NY*.

[15] The Concise Encyclopedia of Statistics. Springer, “Kruskal-Wallis Test,” *New York, NY*, 2008.

[16] W. Haynes, “Wilcoxon Rank Sum Test,” *Encyclopedia of Systems Biology. Springer, New York, NY*.