Driver Vigilance in Automated Vehicles: Investigating Hazard Detection Performance

Eric T. Greenlee, Patricia R. DeLucia, & David C. Newton
Texas Tech University

Introduction

- Drivers of automated vehicles must be able to detect and respond to automation failures, which requires maintaining attention to the roadway for prolonged periods of time.
- Traditional vigilance tasks typically elicit a temporal decline in signal detection performance (Warm et al., 2015).
- Declines in detection performance may lead to the driver not responding to automation failures and potential collisions (NHTSA, 2017).
- Similarities between roadway monitoring during automation and traditional vigilance tasks warrant investigation into drivers' ability to maintain vigilance for prolonged periods of time.
- The current study explored the possibility that a vigilance decrement would present itself in the form of a decline in drivers' ability to detect roadway hazards.
- The researchers hypothesized the following:
  1. Across a 40-minute drive, rate of correct detections of roadway hazards would decline.
  2. Following the drive, subjective workload and task-induced stress would be high.

Experiment

Purpose

- To determine if a vigilance decrement occurs in drivers of automated vehicles who are tasked with detecting roadway hazards

Apparatus

- STISIM Drive™ driving simulation software
- Three 23" LED monitors; 135° horizontal field of view
- Logitech G27 steering wheel and pedals

Procedure

- Participants (N = 11) "drove" a vehicle that automatically maintained 60 mph on a straight, one-lane road and maintained its position in the center of the lane.
- Participants were instructed to monitor the roadway for vehicles that appeared every two seconds out of dense fog and disappeared after 200 ms.
- The majority (95%) of vehicles were stopped safely (i.e., not intruding into the main road), and 5% of the vehicles were stopped unsafely (i.e., intruding 1.25 feet into the main road).
- Participants were instructed to press a button mounted on the steering wheel upon seeing an unsafely-stopped vehicle (critical signal).
- There were four parts to the study:
  1. Pre-Test Short Stress State Questionnaire (Helton, 2004)
  2. Practice drive (5 minutes)
  3. Full drive (40 minutes)
  4. NASA-TLX (Hart & Staveland, 1988) and Post-Test Short Stress State Questionnaire

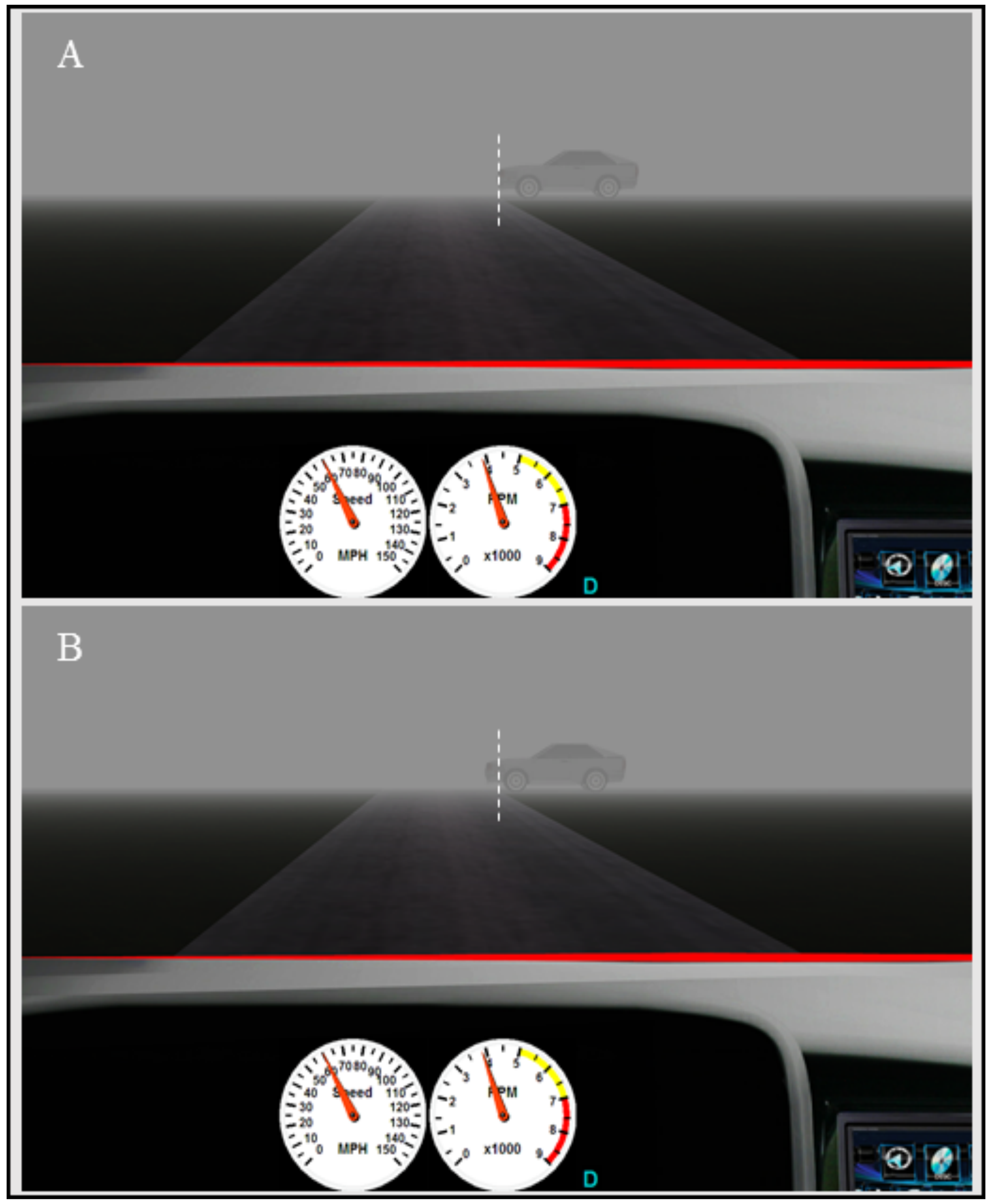
References

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

Helton, W. S. (2004). Validation of a short stress state questionnaire. *Proceedings of the Human Factors and Ergonomics Society, 48,* 1238-1242.

National Highway Traffic Safety Administration (2017). *(Office of Defects Investigation No. PE 16-007)*.

Warm, J. S., Finomore, V. S., Vidulich, M. A., & Funke, M. E. (2015). Vigilance: A perceptual challenge. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman, & J. L. Szalma (Eds.), *The Cambridge handbook of applied perception research,* 241-283. New York, NY: Cambridge University Press.

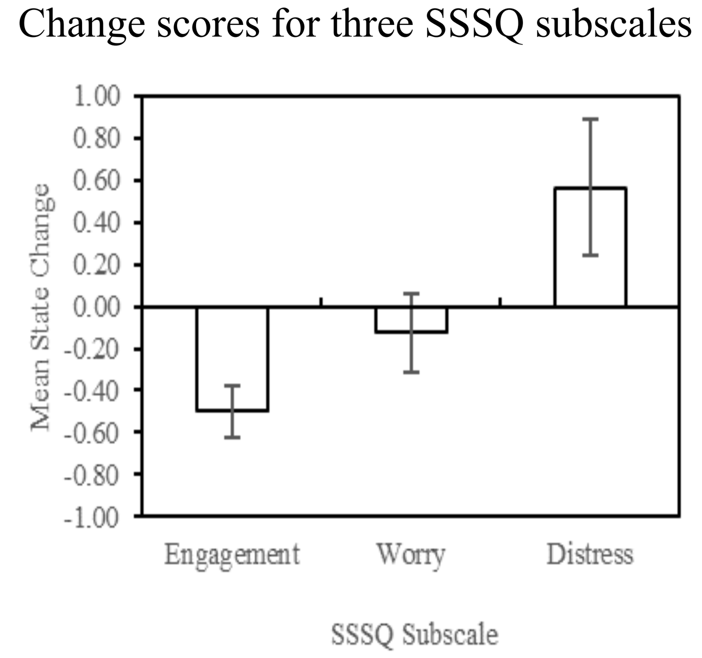


"A" = neutral (safe) stimuli; "B" = critical (unsafe) signals. Dotted vertical lines were not shown.

Results: Stress (SSSQ)

- Change scores for each subscale were analyzed using one-sample *t*-tests (vs. zero).
- Short Stress State Questionnaire change scores indicated a significant decrease in Engagement, $t(9) = 3.96$, $p = .01$, $d = 1.25$. Neither Worry nor Distress changed significantly, $p > .05$.

Change scores for three SSSQ subscales



Note. Error bars are standard errors.

Results: Workload (NASA-TLX)

- A single-factor ANOVA revealed a significant main effect, $F(2.96, 26.65) = 17.08$, $p > .001$, $\eta_p^2 = .66$.
- Mean scores on NASA-TLX subscales of Mental Demand, Temporal Demand, Effort, and Frustration indicated substantial levels of workload (i.e., score > 50).

Mean ratings for NASA-TLX.

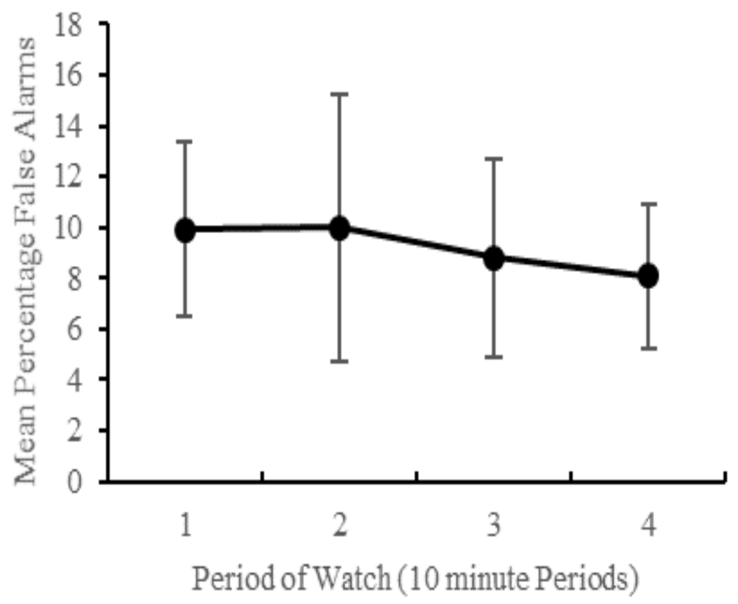
Mental D.	Physical D.	Temporal D.	Effort	Frustration	Performance
77.00	19.40	64.80	72.90	71.40	40.10
(6.98)	(6.13)	(8.09)	(6.34)	(6.45)	(7.84)

Note. Standard errors appear in parentheses. "D" = Demand.

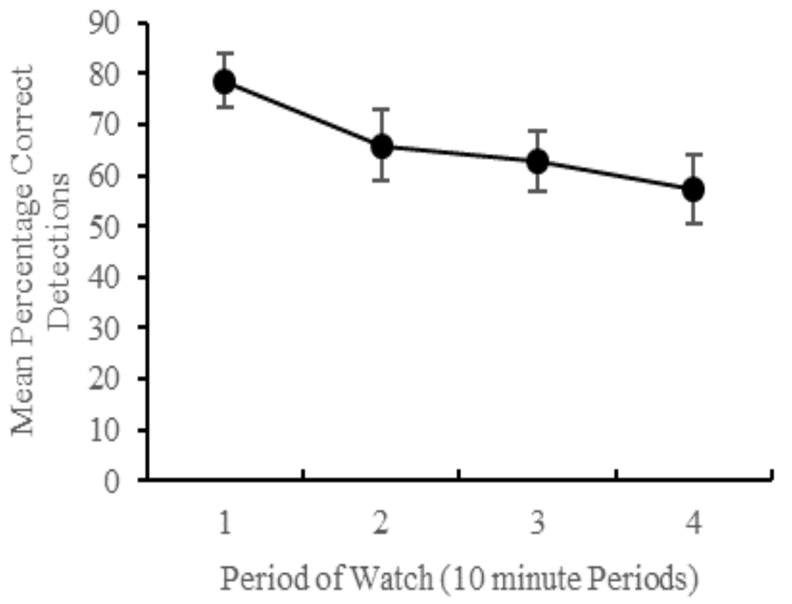
Results: Hazard Detection Performance

- Performance was evaluated with measures of percentage correct detections and percentage false alarms. The 40-minute drive was divided into four 10-minute periods of watch for analyses.
- Mean correct detections declined significantly across the vigil, $F(2.40, 21.63) = 5.55$, $p = .008$, $\eta_p^2 = .38$. Mean correct detections was significantly greater in period 1 than in period 4, $p = .006$.
- Mean false alarms did not differ significantly across the drive, $p > .05$.

Percentage false alarms across drive



Percentage correct detections across drive



Note. Error bars represent +/- 1 standard error of the mean.

Conclusions

- Declines in mean correct detections indicate that the automated driving task induced a vigilance decrement.
- NASA-TLX scores indicate that the task induced high subjective workload and led to decreased task engagement, effects that characterize traditional vigilance tasks.
- Monitoring for potential roadway hazards over prolonged periods may result in a decline in detection performance, high workload, and disengagement which—in turn—may lead to collisions.
- The present findings warrant the need for additional research in order to further understand this phenomenon.