80 MPH and out-of-the-loop: Effects of real-world semi-automated driving on driver workload and arousal.

Francesco N. Biondi, Monika Lohani, Rachel Hopman, Sydney Mills, Joel M. Cooper, David L. Strayer
University of Utah
francesco.biondi@utah.edu

The introduction of semi-automated driving systems is expected to mitigate the safety consequences of human error. Observational findings suggest that relinquishing control of vehicle operational control to assistance systems might diminish driver engagement in the driving task, by reducing levels of arousal. In this study, drivers drove a Tesla Model S with Autopilot in both semi-automated and manual modes. Driver behavior was monitored using a combination of physiological and behavioral measures. Compared to manual driving, a reduction in driver physiological activation was observed during semi-automated driving. Also, performance to the peripheral detection task suffered in semi-automated mode, with slower response times recorded in this condition than during manual driving. Taken together, our data suggest that semi-automated driving might not ease safety consequences of human error. Instead, these findings suggest it might cause a drop in driver monitoring, possibly followed by a spike in automation-generated distraction.

INTRODUCTION

Automated vehicle systems hold the promise of enhancing traffic efficiency, reducing the potential for human error, and making driving safer. In fact, recent estimates suggest that a full adoption of driver assistance technology would reduce the number of road crashes and fatalities by 25% to 90% (European Road Observatory, 2016; Litman, 2017).

The Society of Automotive Engineers defines six levels of vehicle automation, from level-0 or fully-manual driving, to level-5 or fully-automated driving (SAE, 2014). Current semi-automated, level-2 vehicles are equipped with lateral and longitudinal control systems that automatically keep the vehicle within the lane, and maintain it at a certain speed and distance from the forward vehicle, respectively. One requirement of semi-automated driving is for drivers to monitor the functioning of the systems and safely regain control of the vehicle in case of system failures (SAE, 2014).

By stipulating an inverted U-shape relationship between arousal and performance levels, the Yerkes-Dodson law (Yerkes & Dodson, 1908) suggests that extremely low (or extremely high) levels of arousal or stimulation will lead to a decline in human performance. Within the context of driving, the Yerkes-Dodson law is adopted to account for passive fatigue and boredom (Matthews & Davies, 2001). As drivers reach optimal levels of performance, the increase in workload may then lead to an increase in self-regulatory behaviors (Strayer & Fisher, 2015), active distraction and overload (Coughlin, Reimer, & Mehler, 2011).

Studies of driver interaction with advanced driver assistance systems show the potential for vehicle automation to have unintended consequences on situation awareness. In the study by Stanton and Young (2005), for instance, drivers drove a simulated vehicle with or without Adaptive Cruise Control (ACC) engaged. The system maintained the vehicle at a certain speed and distance from the forward car. The authors found that when ACC was engaged this had the potential to reduce levels of driver situation awareness Similar findings were obtained in the study by Vollrath, Schleicher and Gelau (2011), with participants driving a simulated vehicle with increasing levels of longitudinal vehicle automation: fully-manual, Cruise Control (the system maintains the vehicle at a certain speed), Adaptive Cruise Control. Compared to manual driving, using Cruise Control and ACC was associated with a reduction in the number of speed violations. However, further analysis revealed diminished situation awareness when driving with Cruise Control and ACC activated, resulting in drivers' impaired ability to readily respond to upcoming traffic hazards.

What found in the studies by Stanton and Young (2005) and Vollrath et al. (2011) is consistent with what hypothesized by the Yerkes-Dodson Law. As participants relinquished control of vehicle operations to assistance systems, this resulted in reduced driver engagement in the primary task, and, in turn, diminished awareness of the surrounding traffic scenario. Observational data on driver interaction with partial automation confirm this. In her evaluation of a Tesla Model S and its Autopilot system, for instance, Endsley (2017) identified some issues associated with the design and user interface of the semi-automated driving system. First of all, given the limited amount of information provided about the system capabilities and limitations, this caused the driver to develop an inaccurate mental model of the system, an aspect often associated with automation misuse (Parasuraman & Riley, 1997). The author also noticed that the functioning of the *apparently* reliable system, led to episodes of mind-wandering and disengagement from the driving and monitoring task. In turn, this caused delayed responses to unexpected traffic events. In the study by Banks and colleagues (2018) with participants driving an on-road semi-automated vehicle, the authors noticed similar driver disengagement episodes. In particular, drivers appeared to reduce their level of monitoring of the driving task, and, gradually, fall *out-of-the-loop*. For additional findings, see Biondi, Goethe, Cooper and Strayer (2016).

Taken together, these observational findings suggest an indirect association between vehicle automation and driver engagement in the driving task. However, a systematic investigation of the effect of semi-automated driving on the driver state and levels of arousal is missing.

In this study, we adopt a combination of behavioral and physiological measures to examine the potential for on-road semi-automated driving to cause a decline in levels of driver arousal and workload. While driving a Tesla vehicle in Autopilot mode, driver physiological activation was monitored by recording variations in heart rate and heart rate variability. Driver performance to the peripheral detection task was also measured, as slower response times are often observed as a result of under-arousal and sleepiness (Gershon, Shinar, & Ronen, 2009). Self-reported measures of mind-wandering were also collected.

METHOD

Participants

Twenty-two participants (14 males, 8 females) between the age of 21 and 35 (M=25.4) were recruited to participate in this study. All participants had normal or corrected-to-normal vision, did not use cochlear implants or any other hearing device and did not report having hearing deficits. Participants completed a University of Utah IRB-approved consent document and a general demographics survey.

Materials and measures

Vehicle. The Tesla Model S was equipped with Autopilot and Autosteer (version 8.1). These two systems working jointly maintain the vehicle at a certain speed and at the center of the lane. When following a forward vehicle, Autopilot also maintains a set distance to it. The Tesla was also equipped with an Automatic Lane Change system that automatically moves the vehicle to an adjacent lane when instructed by the driver.

Physiological measures. Driver physiological activation was recorded using eMotion Faros 180 heart and respiration rate monitor from Biomation at 250 Hz sampling rate. To record electrocardiogram (ECG) participants wore three electrodes: one placed below the left and right collarbones and the third on the last left rib. Data were collected and analyzed using the Cardiscope software (HASIBA Medican GmBH, Biomation). Spectral electroencephalogram (EEG) was also recorded throughout the study, but would not be reported here.

Continuous ECG data were processed using a custom software (Cardiscope Analytics), to detect R-wave peaks with recommended settings (Berntson, Quigley, & Lozano, 2007; Task Force, 1996). Adopting standard practices, the automatically detected R-wave peaks were visually inspected for accurate detection and manually corrected if improbable values were marked by custom software. After data cleaning, data were processed to calculate the mean heart rate and heart rate variability measures for each phase (manual and semi-automated driving) for each participant. Heart rate was defined as the number of heart beats per minute (Berntson et al., 2007; Task Force, 1996). A higher heart rate is indicative of greater physiological activation and in driving literature is associated with higher workload (Mehler, Reimer, Coughlin, & Dusek 2009). A time domain index of heart rate variability, *triangular interpolation of NN (ms) interval histogram* (TINN) was also estimated. TINN is the baseline width of the distribution measured as a base of a triangle, approximating the normal-to-normal (NN) interval distribution (the minimum square difference is used to find such a triangle; for details see Task Force, 1996). Lower TINN implies lower heart rate variability. TINN was selected as it is a sensitive measure of workload and mental stress (Heine et al., 2017; Orsila et al., 2015).

Peripheral detection task. For the peripheral detection task, participants wore a vibrotactile motor on their left arm, and responded to the onset of the vibration by pressing a micro-switch located on their left index finger against the steering wheel. The motor was placed on the arm to minimize interference with the heart rate recording equipment. The vibration had a duration of 1 second. To minimize the risk of vibrotactile stimuli interfering with natural fluctuations in arousal, we adopted an inter-stimulus interval of 5 minutes.

Subjective measure. During the experimental drives (manual and semi-automated), the presentation of the vibrotactile stimulus also prompted participants to indicate their mind-wandering state. Based upon the protocol developed in the study by Smallwood, Beach, Schooler and Handy (2008), a 3-point scale with the following definitions was developed: (1) fully attentive to the driving task, (2) aware of my thoughts being away from the driving task, (3) immersed in my own thoughts and unaware of this.

Design and Procedure

Design. Two within-subject experimental conditions were implemented: manual and semi-automated driving modes. During manual driving, participants were instructed to drive a 2017 Tesla Model S 70 in manual mode. During semi-automated driving, participants drove the Tesla with Autopilot and Autosteer automation engaged. The order of conditions (manual and semi-automated mode) was counterbalanced across participants.

Procedure. Set-up occurred on the university campus in Salt Lake City (UT). After providing an overview of the study, the heart rate and EEG sensors were placed on the participant following standard protocols (Berntson et al., 2007; Task Force, 1996). Participants were then placed in the driver's seat and the physiological signals were monitored for stability. Next, baseline heart rate and EEG signals were recorded for 5 minutes while parked.

Participants were instructed on how to respond to the vibrotactile stimulus and completed some practice trails until they were comfortable with the task. Participants were then instructed on the 3 levels of mind wandering (Smallwood et al., 2008): 1 = *fully attentive*, 2 = *away from the driving task*, (3) *fully immersed in my own thoughts*. Participants were informed that the vibrotactile motor would vibrate at random intervals as a cue to report their state of mind. They were asked to verbally categorize their state of mind (out of the 3 options) immediately before the motor vibrated as accurately as possible. Instructions were also given on how to activate, use, and adjust Autopilot and Autosteer features during the semi-automated driving phase of the study.

Driving ensued once participants understood all the instructions. Participants drove on Interstate-80 to Aragonite (UT) from Salt Lake City (UT) in either manual or semi-automated mode, and then switched conditions for the return drive. A research associate sat in the rear seat during the entire duration of the drive, and driver behavior was recorded via audio and video recordings. The route was a 63-mile long, four-lane, two-carriageway road with and an annual average daily traffic of about 9,000 vehicles (UDOT, 2018). The speed limit varies from 70 MPH to 80 MPH along the route. In between the two drives, participants were given a 15-minute break at a rest stop in Aragonite (UT). During each drive, performance to the peripheral detection task, state of mind, and continuous physiological data were collected. After completing each mode, participants also reported subjective ratings of workload. Data analysis was conducted in R (R Core Team, 2000).

RESULTS

Peripheral detection task. Data from four participants were excluded from the analysis due to recording issues. Response times collected for the peripheral detection task were analyzed. Given the low number of misses (accuracy > 99.9%), accuracy data were not analyzed. A paired t-test was conducted to investigate whether RT collected during semi-automated driving differed from those collected during manual driving. Participants responded significantly slower when driving in semi-autonomous mode (M=1068, SE=41.87) compared to manual mode (M=879, SE=36.16), as shown through a t-test, $t(17) = 2.69$, $p<.05$).

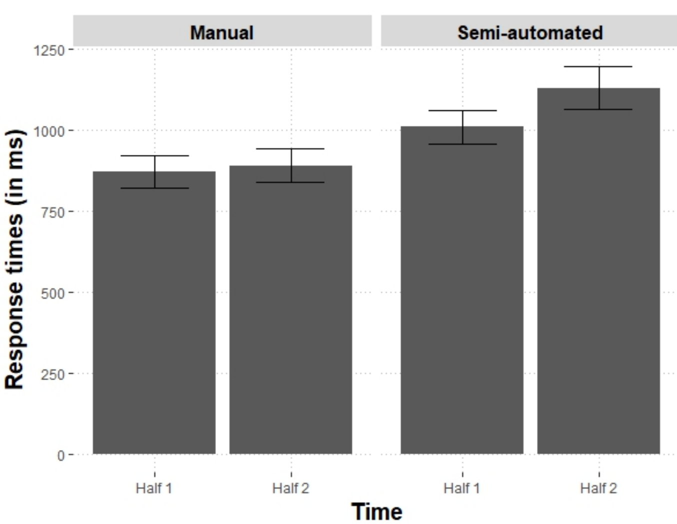

Figure 1. Average response times (in ms) to the peripheral detection task by driving mode (manual vs. semi-automated) and time of the drive (Half 1 and 2). Error bars represent standard error.

Physiological measures. Physiological data from ten participants were removed from the analyses due to recording and artifact issues during one or two drives. A t-test with Mode (2 levels: semi-automated, manual) as independent variable and average heart rate as dependent variable revealed lower heart rate during semi-automated driving (M = 72.12 bpm, SE = 2.62), compared to manual driving (M = 75.03 bpm, SE = 2.68), $t(11) = 3.28$, $p < .05$.

Heart rate variability calculated as TINN was also compared across the two driving modes. Relative to manual driving (M = 413.33 ms, SE = 40.45), driving the vehicle in semi-automated mode led to higher heart rate variability (M = 468 ms, SE = 36.36). This suggests an increase in parasympathetic activity during semi-automated driving, compared to manual driving. No differences in respiration rate were found: manual (M = 18.56 rate per minute, SE = .44) vs. automated (M = 18.93 rate per minute, SE = 2.22), p = .15. This indicates that the differences in heart rate variability were not caused by task differences in respiration rate in manual vs. semi-automated driving modes. Average heart rate variability is presented in Figure 2.

Self-reported measure. Pearson's Chi-squared analysis did not reveal significant differences between semi-automated and manual driving in self-reported mind-wandering (on the 3-point scale), $X^2(2) = 1.33$, $p > .05$.

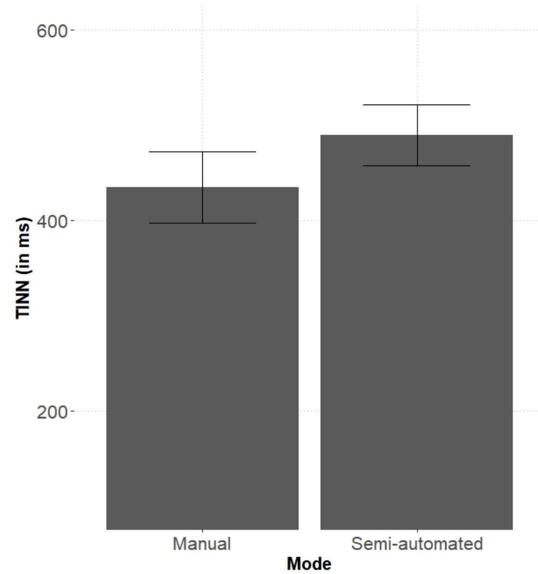

Figure 2. Heart rate variability calculated as triangular interpolation of NN or TINN in ms. Data are presented by mode (manual vs. semi-automated). Error bars represent standard error.

DISCUSSIONS

Levels of driver arousal decreased during semi-automated driving, and their responsiveness to peripheral stimuli diminished as a result of driving in Autopilot mode. Compared to manual driving, average heart rate decreased and heart rate variability increased when driving with assistance systems activated. This pattern of results is consistent with the literature on mind-wandering, showing an increase in parasympathetic activity when transitioning into a state of relaxation (Ottaviani, Shapiro, & Couyoumdjianb, 2013).

Studies on the effect of workload on physiological activation show a direct association between average heart rate and cognitive load. In the study by Mehler, Reimer, Coughlin, & Dusek (2009), for instance, an increase in average heart rate from 73.6 bpm to 79.4 bpm was recorded with drivers completing more cognitively demanding versions of the n-back task (0-back to 2-back). Similarly, in the study by Biondi, Coleman, Cooper and Strayer (2016), drivers experienced higher heart rate as the secondary-task became more demanding. In this study, the reduction in average heart rate and increase in heart rate variability recorded during semi-automated driving is then indicative of the reduced physiological activation experienced by drivers as a result of their relinquishing vehicle operations to assistance systems.

Performance to the peripheral detection task is also sensitive to mental load, or the lack thereof. In the study by Gershon, Shinar and Ronen (2009), participants drove a simulated vehicle for 2 hours, while responding to the onset of one of three colored lights that appeared on the dashboard, in the vicinity of the steering wheel. As participants became more fatigued (an increase in subjective measures of fatigue and heart rate variability was observed over time), performance to the detection task suffered and response times became slower. Studies on the effect of overload on the performance in the Detection Response Task or DRT (ISO, 2016) show consistent results, with longer response times and lower accuracy recorded as the amount of cognitive resources available to drivers decreases. (Harbluck, Burns, Hernandez, & Glazduri, 2013; Jenness et al., 2016). A pattern of results consistent with that of Gershon et al. (2009) was found in our study. The drop in mental workload that characterized semi-automated driving, resulted in slower response times in the peripheral detection task. In particular, response times recorded in Autopilot mode were longer than those recorded during manual driving by about 188ms, a greater difference than that typically found in distraction studies (Strayer et al., 2015).

One of the main risks associated with vehicle automation is for drivers to gradually become under-stimulated and, eventually, fall *out-of-the-loop* (Cunningham, & Regan, 2015). Research in user interaction with automation in the aviation and maritime fields suggest that human performance in monitoring tasks deteriorates over prolonged amount of times, as a result of the increase in complacency toward the system and/or the fragmented knowledge toward its capabilities and limitations (Bailey & Scerbo, 2007).

Surprisingly, levels of self-reported mind-wandering were not affected by driving mode. Contrary to our expectations, subjective data did not seem to reflect a decline in driver engagement in the monitoring task as a result of driving in semi-automated mode. A plausible explanation to this can be found in the literature on the effect of cognitive overload on self-awareness. In the study by Sanbonmatsu and colleagues (2015), participants drove a simulated vehicle in manual mode while being engaged in a cell phone task. Relative to the control condition with no secondary task, the cognitive overload originated from the concurrent cell phone conversation, and subsequent reduction in mental resources directed to driving, resulted in a decline in the driver performance, and, interestingly, their self-awareness of that impairment. In other words, overloaded drivers were not only worse at driving, but they had a poor awareness of that decline in performance. In our study, we hypothesize the condition of underload experienced during semi-automated driving to have impaired driver self-awareness, and, in turn, their ability to accurately estimate their state.

Our data document the direct consequences of semi-automated driving on under-arousal. However, further analysis is needed in order to investigate temporal fluctuation in driver physiological activation, and to examine how increased exposure to semi-automated driving would affect heart rate and heart rate variability. As drivers transition from being novices to experts, we predict them to become even more disengaged from driving, and possibly show signs of self-regulation and active distraction.

This study demonstrates the potential for semi-automated driving to have unexpected safety consequences. Our findings are consistent with the literature on user automation, and help shed light on the causes of the recent accidents involving semi-automated driving (NTSB, 2017). Our participants had no prior experience with level-2 automated driving systems, were monitored during the data collection, and used Autopilot for a limited amount of time. Despite this, drivers showed signs of under-arousal during semi-automated driving (videos show participants yawning, or having difficulty keeping their eyes open), resulting in reduced vigilance and awareness of traffic hazards. Future work will compare novice and expert Tesla drivers as we predict a greater drop in arousal to produce an increase in self-regulatory activities for experts.

Acknowledgments. We thank the University of Utah Research Foundation, and the AAA Foundation for Traffic Research for their support to this research project.

REFERENCES

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, *68*(October 2017), 138–145. http://doi.org/10.1016/j.apergo.2017.11.010

Bailey, N.R., Scerbo, M.W., 2007. Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science* 8 (4), 321–348. https://doi.org/10.1080/14639220500535301

Berntson, G. G., Quigley, K. S., & Lozano, D. (2007). Cardiovascular psychophysiology. *Handbook of psychophysiology*, *3*, 182-210.

Biondi, F., Goethe, R., Cooper, J., & Strayer, D. (2017). Partial-autonomous Frenzy: Driving a Level-2 Vehicle on the Open Road. In *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 329–338). Springer, Cham.

Coughlin, J. F., Reimer, B., & Mehler, B. (2011). Monitoring, managing, and motivating driver safety and well-being. *IEEE Pervasive Computing*, *10*(3), 14–21. http://doi.org/10.1109/MPRV.2011.54

Cunningham, M., & Regan, M. A. (2015). Autonomous Vehicles: Human Factors Issues and Future Research. *Proceedings of the 2015 Australasian Road Safety Conference.* 14 - 16 October, Gold Coast, Australia

Endsley, M. R. (2017). Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, *11*(3), 225–238. http://doi.org/10.1177/1555343417695197

European Road Safety Observatory. (2016). *Advanced driver assistance systems*. Retrieved from https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/ersosynthesis2016-adas15_en.pdf

Gershon, P., Shinar, D., & Ronen, A. (2009). Evaluation of experience-based fatigue countermeasures. *Accident Analysis and Prevention*, *41*(5), 969–975. http://doi.org/10.1016/j.aap.2009.05.012

Harbluk, J. L., Burns, P. C., Hernandez, S., Tam, J., & Glazduri, V. (2013). Detection response tasks: using remote, headmounted and tactile signals to assess cognitive demand while driving. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 78–84.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North Holland.

Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., & Deml, B. (2017). Electrocardiographic features for the measurement of drivers' mental workload. *Applied Ergonomics*, *61*, 31–43. http://doi.org/10.1016/j.apergo.2016.12.015

International Organization for Standardization. (2016). *ISO 17488:2016. Road vehicles — Transport information and control systems — Detection-Response Task (DRT) for assessing attentional effects of cognitive load in driving*. Retrieved from https://www.iso.org/standard/59887.html

Jenness, J. W., Boyle, L. N., Lee, J. D., Chang, C., Venkatraman, V., Gibson, M., … Kellman, D. (2016). In-Vehicle Voice Control Interface Performance Evaluation (Final report. Report No. DOT HS 812 314). *National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT)*, (October).

Litman, T. (2017). Autonomous Vehicle Implementation Predictions. Implications for Transport Planning. *Victoria Transport Policy Institute*, *28*. Retrieved from https://www.vtpi.org/avip.pdf

Matthews, G., & Davies, D. R. (2001). Individual differences in energetic arousal and sustained attention: a dual-task study. *Personality and Individual Differences, 31*, 575–589. https://doi.org/10.1016/S0191-8869(00)00162-8

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J(2009). Impact of Incremental Increases in Cognitive Workload on Physiological Arousal and Performance in Young Adult Drivers. *Transportation Research Record: Journal of the Transportation Research Board*, (2138), 6–12. http://doi.org/10.3141/2138-02

NTSB. (2017). Collision between a Car Operating with Automated Vehicle Control Systems and a Tractor-Semitrailer Truck, Williston, FL, May 7, 2016. Retrieved from https://www.ntsb.gov/news/events/Documents/2017-HWY16FH018-BMG-abstract.pdf

Ottaviani, C., Shapiro, D., & Couyoumdjian, A. (2013). Flexibility as the key for somatic health: From mind wandering to perseverative cognition. *Biological Psychology*, *94*(1), 38–43. http://doi.org/10.1016/j.biopsycho.2013.05.003

Orsila, R., Virtanen, M., Luukkaala, T., Tarvainen, M., Karjalainen, P., Viik, J., & Savinainen, M. (2008). Perceived mental stress and reactions in heart rate variability—a pilot study among employees of an electronics company. *International Journal of Occupational Safety and Ergonomics*, *14*(3), 275–283. http://doi.org/10.1080/10803548.2008.11076767

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. http://doi.org/10.1518/001872097778543886

R Core Team (2000). R language definition. *Vienna, Austria: R foundation for statistical computing*.

SAE (2014). *Automated Driving. Levels of Driving Automation Are Defined in New SAE Internaional Standsd J3016*. Retrieved from http://www.sae.org/misc/pdfs/automated_driving.pdf

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the Brain: Mind Wandering Reduces Cortical Analysis of External Events. *Journal of Cognitive Neuroscience*, *20*(3), 458–469. http://doi.org/10.1162/jocn.2008.20037

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, *48*(10), 1294–1313. http://doi.org/10.1080/00140130500252990

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing Cognitive Distraction in the Automobile. *Human Factors*, *57*(8), 1300–24. http://doi.org/10.1177/0018720815575149

Strayer, D. L., & Fisher, D. L. (2015). SPIDER: A Framework for Understanding Driver Distraction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720815619074-. http://doi.org/10.1177/0018720815619074

Task Force (1996). Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, *93*(5), 1043-1065.

UDOT (2018). Traffic Statistics. Utah Department of Transportation. Retrieved from https://www.udot.utah.gov/main/f?p=100:pg:0:::1:T,V:507

Vollrath, M., Schleicher, S., & Gelau, C. (2011). The influence of Cruise Control and Adaptive Cruise Control on driving behaviour - A driving simulator study. *Accident Analysis and Prevention*, *43*(3), 1134–1139. http://doi.org/10.1016/j.aap.2010.12.023

Yerkes, R. M., and Dodson, J. D. (1908) The relation of strength of stimulus to rapidity of habit- formation. *Journal of Comparative Neurology of Psychology*, 18, 459-482.