No Difference in Arousal or Cognitive Demands Between Manual and Partially Automated Driving: A Multi-Method On-Road Study

Monika Lohani[1]*, Joel M. Cooper[2], Gus G. Erickson[2], Trent G. Simmons[2], Amy S. McDonnell[2], Amanda E. Carriero[2], Kaedyn W. Crabtree[2] and David L. Strayer[2]

[1] Department of Educational Psychology, University of Utah, Salt Lake City, UT, United States, [2] Department of Psychology, University of Utah, Salt Lake City, UT, United States

Introduction: Partial driving automation is not always reliable and requires that drivers maintain readiness to take over control and manually operate the vehicle. Little is known about differences in drivers' arousal and cognitive demands under partial automation and how it may make it difficult for drivers to transition from automated to manual modes. This research examined whether there are differences in drivers' arousal and cognitive demands during manual versus partial automation driving.

Method: We compared arousal (using heart rate) and cognitive demands (using the root mean square of successive differences in normal heartbeats; RMSSD, and Detection Response Task; DRT) while 39 younger (M = 28.82 years) and 32 late-middle-aged (M = 52.72 years) participants drove four partially automated vehicles (Cadillac, Nissan Rogue, Tesla, and Volvo) on interstate highways. If compared to manual driving, drivers' arousal and cognitive demands were different under partial automation, then corresponding differences in heart rate, RMSSD, and DRT would be expected. Alternatively, if drivers' arousal and cognitive demands were similar in manual and partially automated driving, no difference in the two driving modes would be expected.

Results: Results suggest no significant differences in heart rate, RMSSD, or DRT reaction time performance between manual and partially automated modes of driving for either younger or late-middle-aged adults across the four test vehicles. A Bayes Factor analysis suggested that heart rate, RMSSD, and DRT data showed extreme evidence in favor of the null hypothesis.

Conclusion: This novel study conducted on real roads with a representative sample provides important evidence of no difference in arousal and cognitive demands. Younger and late-middle-aged motorists who are new to partial automation are able to maintain arousal and cognitive demands comparable to manual driving while using the partially automated technology. Drivers who are more experienced with partially automated technology may respond differently than those with limited prior experience.

Keywords: heart rate, heart rate variability, detection response task, partial driving automation, applied cognition

INTRODUCTION

Some commercially available vehicles with partial vehicle automation can support safe driving (e.g., Tesla-Autopilot, Nissan-ProPilot, Volvo-Pilot Assist, and Cadillac-Supercruise). However, partial vehicle automation is not always reliable, requiring drivers to maintain readiness to take over vehicle control at all times (SAE, 2016). If drivers' cognitive demands during partially automated driving are different from manual driving mode, it may raise concerns about drivers' cognitive readiness to take over, should automation fail. In studying cognitive demands during manual and partially automated driving, it is important to consider driver's arousal (Yerkes and Dodson, 1908; Hebb, 1955; Broadhurst, 1959; Wekselblatt and Niell, 2015). The concern with partially automated technology is that it may lead to suboptimal arousal levels and cognitive demands resulting in poor driving performance. This motivated the study's research question: Are there differences in drivers' arousal and cognitive demands during manual versus partially automated driving? A comparison of physiological arousal and cognitive demands during the two modes of driving would help understand potential differences (or lack thereof) in cognitive demands and subsequent driving performance. The aim of the current study was to investigate whether there are differences in drivers' arousal and cognitive demands during manual versus partial automation driving on real highways.

Only limited research is currently available to understand potential differences in drivers' cognitive demands under partial vehicle automation. Based on classic cognitive models of attention (Kahneman, 1973; Wickens, 2002), *driving-related cognitive demands* are defined as all cognitive and mental processing resources required to perform a driving task. Some studies have suggested that there may be significant cognitive demands (e.g., workload) during automated driving compared to manual driving (Solís-Marcos et al., 2018; Kim et al., 2020). However, other research has found cognitive demands during partially automated driving may actually be reduced (Biondi et al., 2018; Heikoop et al., 2019; Zhai and Lu, 2019). By contrast, other research suggests little or no difference (Sibi et al., 2017; Stapel et al., 2019; Calvi et al., 2020; Várhelyi et al., 2020). Note that Stapel et al. (2019) found no difference in automation for inexperienced drivers; however, lower workload was found with experienced drivers. A variety of factors can help explain these inconsistent findings, including small sample sizes limiting the likelihood of detecting a true effect; testing only a single vehicle raising questions of ecological validity; and use of self-reports administered after driving raising concerns of retrospective report biases. Furthermore, most of the research on this topic has been restricted to samples of younger adults that limits the applicability of findings to the general public. Thus, a clear understanding of whether and how partially automated technology influences drivers' arousal and cognitive demands are still lacking.

Arousal is a heterogeneous construct that involves general autonomic activation (e.g., Robbins and Everitt, 1995; Satpute et al., 2019). Heart rate is a biomarker of arousal and they are positively related (Berntson et al., 2007; Mauss and Robinson,

2010; Satpute et al., 2019). Heart rate is commonly used to measure driving-related arousal changes (Lohani et al., 2019) and it was used to operationalize arousal in the current study. Moreover, in real-world driving (for a review see, Lohani et al., 2019), multiple overlapping constructs dynamically change simultaneously and interdependently (e.g., workload, stress, boredom, distraction) resulting in net cognitive demands experienced by a motorist that can be broadly categorized along a low-to-high spectrum. Low cognitive demands may be represented by a combination of constructs such as drowsiness and boredom, while high cognitive demands may be represented by mental workload and stress (Lohani et al., 2019). Cognitive demands are associated with *cardiac vagal control*, which represents the influence of the vagus nerve on heart functioning (for a review, see Berntson et al., 2007; Laborde et al., 2017). Cardiac vagal control can be indexed by vagally-mediated *heart rate variability* (HRV), i.e., the temporal variability in adjacent heartbeats (Malik, 1996; Berntson et al., 2007). According to the neurovisceral integration model, the neural circuitry for cognitive and autonomic regulation has an overlapping neurovisceral mechanism (see Thayer et al., 2009; Smith et al., 2017), which can explain the coupling between cognitive demands and vagally-mediated HRV. Thus, vagally-mediated HRV can measure changes in cognitive demands, such as mental effort, workload, and attention (e.g., Mulder and Mulder, 1981; Mulder, 1985; Thayer et al., 2009).

Indeed, vagally-mediated HRV has been shown to provide a near-real-time objective measure of dynamic changes in cognitive demands, making it a suitable measure for applied driving research without disrupting the driving task (e.g., Lee et al., 2007; Liang et al., 2009; Mehler et al., 2011; Noda et al., 2015; Sugie et al., 2016; Heine et al., 2017; Piotrowski and Szypulska, 2017; Cisler et al., 2019; Lohani et al., 2019). Some of these studies found that cognitive state detection while driving was better when vagally-mediated HRV was utilized in addition to behavioral measures (e.g., Lenneman and Backs, 2009). Even though a variety of measures of vagally-mediated HRV have been used in applied research, the root mean square of successive differences in normal heartbeats (RMSSD) has been found to change more systematically with driving-related cognitive demands (Mehler et al., 2011; Heine et al., 2017), and thus was a suitable measure for the current study. In particular, RMSSD decreases with an increase in cognitive demands while driving (Mehler et al., 2011) due to the links between vagally-mediated HRV and neural activity associated with cognitive regulation (e.g., Thayer et al., 2009). Furthermore, a standard behavioral method to assess cognitive demands in driving research is the Detection Response Task (DRT; International Organization for Standardization [ISO], 2016). Performance on the DRT (International Organization for Standardization [ISO], 2016) is a measure of driving-related cognitive demands due to visual demands and driving difficulty (Bengler et al., 2012; Bruyas and Dumont, 2013; Cooper et al., 2016). An increase in driving-related cognitive demands is associated with increased reaction time performance and decreased hit-rate on the DRT (Young et al., 2013). At the same time, drowsiness and fatigue also increase reaction time on this task (Gershon et al., 2009). DRT

was included as a behavioral measure of cognitive demands in the current study.

In prior research, we conducted a pilot study with 28 young drivers (M_{age} = 29.29 years) new to partial automation who drove three partially automated vehicles on a flat and straight section of interstate highway (Lohani et al., 2020). In this pilot study, there were no differences between manual and partially automated driving modes across outcomes: heart rate, RMSSD, electroencephalogram (EEG) alpha, and theta power, and DRT performance. The current study is a new follow-up study designed to replicate and extend the earlier findings with a larger and more representative sample of younger and late-middle-aged drivers, partially automated vehicles (Cadillac, Nissan Rogue, Tesla, and Volvo), and sections of roadway. The current research measured heart rate to compare arousal and RMSSD and DRT performance to compare cognitive demands under manual versus partially automated driving.

Based on previous work, the current study considered three alternative hypotheses. First, if partially automated driving leads to high arousal and high cognitive demands (e.g., workload and stress; Solís-Marcos et al., 2018; Kim et al., 2020), then, a significant increase in heart rate, a decrease in RMSSD, and an increase in the DRT reaction time rate would be expected when compared to manual driving. Second, if partially automated driving leads to low arousal and low cognitive demands (e.g., drowsiness and boredom; Biondi et al., 2018; Heikoop et al., 2019; Zhai and Lu, 2019), then a significant decrease in heart rate, an increase in RMSSD, and an increase in DRT reaction time would be expected when compared to manual driving. Finally, if arousal and cognitive demands are similar under manual and partially automated driving, then no differences in the outcome measures would be expected (e.g., Sibi et al., 2017; Stapel et al., 2019; Calvi et al., 2020; Lohani et al., 2020; Várhelyi et al., 2020). This prediction of the null hypothesis has a compelling rationale and a meaningful interpretation that would imply that manual and partially automated modes are comparable in arousal and cognitive demands. However, some limitations (e.g., small sample size, a non-representative sample of people and vehicles, and inadequate statistics) can hamper the ability to interpret evidence for a null hypothesis adequately. Therefore, we designed the current study to allow for a fair test of the null hypothesis by testing a larger and more representative sample of drivers and vehicles. We bolstered the interpretation by conducting the classic null-hypothesis significance testing and a Bayesian alternative (Kruschke, 2011) to meaningfully interpret whether the current dataset supported the null or the alternative hypothesis.

MATERIALS AND METHODS

Participants

A total of 71 adults with no prior experience with partially automated vehicles participated in this study. 39 participants were younger (21–42 age range, M_{age} = 28.82 years, SD_{age} = 6.41, 13 females). 32 participants were late-middle-aged (43–64 age range, M_{age} = 52.72 years, SD_{age} = 6.33, 12 females). The study

protocol was in accordance with the Institutional Review Board at the University of Utah. Participants had no previous experience with partial automation, had a valid driver's license, no at-fault accidents in the past 2 years, drove at least an average of 10 h per month, had no history of neurological disorders or heart conditions, and were not pregnant. In addition, participants were required to pass a 20 min online defensive driving course and certification test. Upon arrival in the lab, eligible participants were allowed to participate in the study only if they had slept at least 6 h the previous night and had their blood alcohol level at 0.0%, which was verified using a BACtrac breathalyzer.

Measures

Psychophysiological data was continuously sampled at 2,000 Hz using a portable wireless physiology system (Smart Center, Biopac System Inc., United States) and Acqknowledge software (Biopac System Inc., United States). This setup allowed real-time data quality monitoring while participants drove on the highway.

Electrocardiography

The electrical activity of the heart was recorded by using an electrocardiogram (ECG). After cleaning the site, standard disposable electrodes were placed on the right collar bone and the left and right end of the ribcage (Lead II configuration; Berntson et al., 2007). During data collection, the ECG was monitored by a research assistant who sat in the front passenger seat. Any noticeable movements that could add artifacts (e.g., sneezes and itch) were marked.

Standardized methods in accordance with recommended guidelines for ECG data were followed (e.g., Malik, 1996; Berntson et al., 1997, 2007; Peltola, 2012; Shaffer et al., 2014; Laborde et al., 2017). Post data collection, the ECG data was processed using AcqKnowledge software (Biopac System Inc., United States). The raw data was bandpass filtered at 1 and 35 Hz cutoffs. The software was used to detect R-wave peaks. All R-wave peaks were visually inspected for accurate detection and manually corrected if the software marked improbable values. This included any artifacts generated by facial or head movements (e.g., yawns or checking blind spots while driving). After data cleaning, data were processed to calculate RMSSD and heart rate for manual and partially automated driving tasks for each vehicle operated by each participant. Based on the recommended guidelines (Malik, 1996; Berntson et al., 1997; Laborde et al., 2017), RMSSD was calculated using the same length epochs (1 min) for the pre-condition baseline and the main-condition periods (see the procedure for details). The epochs were then averaged over the entire period to get average RMSSD values during the pre-condition and main-condition periods. For RMSSD and heart rate data, any values that exceeded three standard deviations from the mean of normal distribution were removed before analyses.

Detection Response Task (DRT)

To probe drivers' cognitive demand, participants were asked to perform a vibrotactile detection task, DRT. In line with the ISO 17488 guidelines (2016), a vibrotactor (a small vibration motor) emitted a small vibration stimulus, similar to a vibrating cell

phone. This stimulus was presented pseudo-randomly every 3–5 s. Participants wore a vibrotractor that was taped to their left biceps. A microswitch was attached to either the index or middle finger of the left hand, which participants could press against the steering wheel to respond to the onset of the stimulus. Instead of the standard left collarbone placement, the left bicep was used to avoid any potential interference with the ECG signal. A similar approach has been successfully used in past work (e.g., Lohani et al., 2020). Participants' goal was to respond to the stimulus onset as quickly as possible while driving (with priority always given to safe driving practices). Response time in milliseconds was recorded for each stimulus. The vibration stimulus was set up to turn off after 1 s.

The average reaction time performance on DRT was calculated for each participant in each vehicle in manual and partially automated driving tasks. Any values that exceeded three standard deviations from the mean of normal distribution were removed before analyses. The hit-rate performance for the current sample was at a ceiling level (∼95% and above). The average proportion of hit-rate during manual driving was 96% ($SE = 0.003$) and during partially automated driving was 95% ($SE = 0.004$). Because of a lack of variance in performance, the hit-rate was not further analyzed.

Self-Reported Experiences With Automation
After driving partially automated vehicles, participants responded to a list of questions about their experiences and attitudes about partially automated vehicles. An example item was "I was anxious and nervous when the automated driving systems were on." Participants indicated their agreement to the following statements using a 5-point scale anchored by completely disagree to completely agree.

Vehicles
We examined a representative sample of commercially available vehicles for this study. A 2018 Cadillac CT6, 2019 Nissan Rogue, 2018 Tesla Model 3, and a 2018 Volvo XC90 were used in this study. Each of these vehicles was equipped with the partial vehicle automation that centered the vehicle within the lane (e.g., Lane Centering) and maintained the following distance and speed (e.g., Adaptive Cruise Control). These features, when activated together, meet the definition of partial automation (SAE, 2016).

Procedure
Participants were sent a training document and a short video about the partial automation features of the vehicle they would drive for the day of the visit. Upon arrival in the lab, after completing the consent form and inclusion criteria testing, participants were set up for ECG data collection. Next, a research assistant directed the participant to the designated vehicle in the parking lot close to the lab. Participants were instructed that they were driving commercially available vehicles, and we were interested in examining the vehicular systems on real roads. They were instructed to operate the vehicle as they would usually drive on the road.

Before driving the vehicle, participants would familiarize themselves with the vehicle and get trained on steps to activate its partial automation features. They engaged the partially automated systems during the training phase, and only when they were comfortable operating the vehicle was the main part of the study started. Participants were asked to keep their hands on the wheel and monitor the road (as recommended by most vehicle manufacturers). In the partial automation condition, participants always had the automation engaged. In rare instances, participants took control of the vehicle to pull over for emergency vehicles, debris on the road, and construction. Participants drove on the same road in manual and partial automation (counterbalanced). As soon as possible, partial automation was re-engaged. It is important to note that participants drove on the same road in both manual and partial automation (counterbalanced), so driving conditions were equated as best as possible. Moreover, any section of the drive where the driver had to deal with an event (e.g., pulling over for an emergency vehicle) were excluded from the analyses (in both manual and partial automation) to ensure a fair comparison.

Participants were also fitted with the DRT equipment and trained on how to perform the DRT task while driving. ECG data were monitored for quality. Next, they drove on a training route with the research assistant in the passenger seat to ensure that the participant could engage and disengage partial automation, change lanes, and control the vehicle's speed. After confirming that participants understood how to use the partial automation and the DRT task while driving safely, the study's testing phase was initiated.

For the testing phase, all participants drove comparable distances on two highways (I-15 and I-80) in two modes (manual and partial automation), with the order partially counterbalanced across participants (leading to 4 experimental sessions in each vehicle). The Average Annual Daily Travel (AADT) for I-15 (N/S bound 4–5 lanes in each direction separated by a median) is 175,000 vehicles and for I-80 (E-W bound 2–3 lanes in each direction separated by a median) is 19,000 vehicles (Utah Department of Transportation). Because there were no significant differences in outcomes between the two highways, the data were averaged for the two highways. The manual and partially automated conditions were also counterbalanced to control for any unsystematic differences.

Each of the four experimental sessions was about 20 min and began once the vehicle was at the posted speed limit. The first 2 min of each session were used as the pre-condition baseline measurement. The following 18 min of the experimental session were the main condition during which the DRT task was performed while participants drove in manual or partial automation conditions (depending on the order). Participants received a short rest break between each of the four experimental sessions. The average value of RMSSD (calculated by averaging 1 min epochs) in the pre-condition baseline period was subtracted from the average value of the main condition to account for any baseline differences within participants. The average heart rate in the pre-condition baseline was subtracted from the main condition. After accounting for baseline differences and collapsing across the two highways, there were two average values for each vehicle driven in manual and partial automation.

Data Analysis Plan

The primary research focus was to examine how automation (manual versus partial automation) affects driver arousal and cognitive demand. To address this question, two analytical approaches were adopted – linear mixed-effects models and Bayes Factor analysis. For each of the outcome measures, a linear mixed-effects model was run with three fixed factors, automation (manual and partially automated), age (younger or late-middle-aged), and vehicle (Cadillac, Nissan Rogue, Tesla, and Volvo) and participants as the random intercept. Note that preliminary analyses were performed to determine if time should be included as a factor in the model. We found that including time as a factor in the model did not explain any significant variability. In contrast to automation condition and age, we had no theoretical or empirical rationale to include time in the models. Thus, to keep a parsimonious model, time was not included in the model because it did not explain significant variability in the outcome variables.

Next, Bayes Factor analysis was conducted to adopt a Bayesian alternative to evaluate how meaningful a significant difference or a lack of significant difference was between the manual and partial automation for each of the outcome variables. This was done by comparing a full model with automation, age, vehicle as predictors to a restricted model without automation (e.g., Kruschke, 2011). The Bayes Factor value is the ratio of the marginal likelihoods of the full model and the restricted model (Lee and Wagenmakers, 2014). The resulting Bayes factor value was interpreted using the classification scheme such that a value higher than 1 is interpreted as evidence in favor of the alternative hypothesis. A Bayes Factor value between 1 and 3 provides anecdotal evidence, 3–10 provides moderate evidence, 10–30 provides strong evidence, 30–100 provides very strong evidence, and a value more than 100 is extreme evidence in favor of the alternative hypothesis. By contrast, a value lower than 1 is interpreted as evidence in favor of the null hypothesis. Correspondingly, a value of 1–0.33 provides anecdotal evidence, 0.33–0.1 provides moderate evidence, 0.1–0.03 provides strong evidence, 0.03–0.01 provides very strong evidence, and a value less than 0.01 provides extreme evidence in favor of the null hypothesis (Lee and Wagenmakers, 2014; Quintana and Williams, 2018). All analyses were done by using the R language for statistical computing (R Core Team, 2020). Mixed models were fit using the *lme4* package (Bates et al., 2015), and

Bayes Factor values were calculated via the *BayesFactor* package (Morey et al., 2018).

RESULTS

See **Table 1** for descriptive statistics for the three outcome variables as a function of age and vehicle. **Figures 1–3** present individual data points to complement the descriptive tables. These include violin plots that are similar to box plots, but in addition, they have the kernel probability density of all the observed data. The mean values are also indicated by a dot in the center of each distribution.

Age differences and gender differences were examined during the pre-condition baseline and main condition. In the pre-condition baseline values of RMSSD, there were significant gender and age-related differences with higher mean values for males ($M = 29.5$, $SE = 1.76$) than females ($M = 22.7$, $SE = 2.27$), $t(69) = 2.38$, $p = 0.02$, and higher RMSSD values for younger ($M = 30.2$, $SE = 1.96$) than late-middle-aged ($M = 21.9$, $SE = 2.10$) drivers, $t(69) = 2.89$, $p = 0.01$. However, after accounting for pre-condition values (average main condition – average pre-condition baseline values), neither gender differences ($p = 0.19$) nor age differences ($p = 0.89$) were significant. We did not have specific hypotheses for gender-related differences in driving manual and partially automated vehicles. Thus, gender was not included as a factor in the models to test the study's research questions.

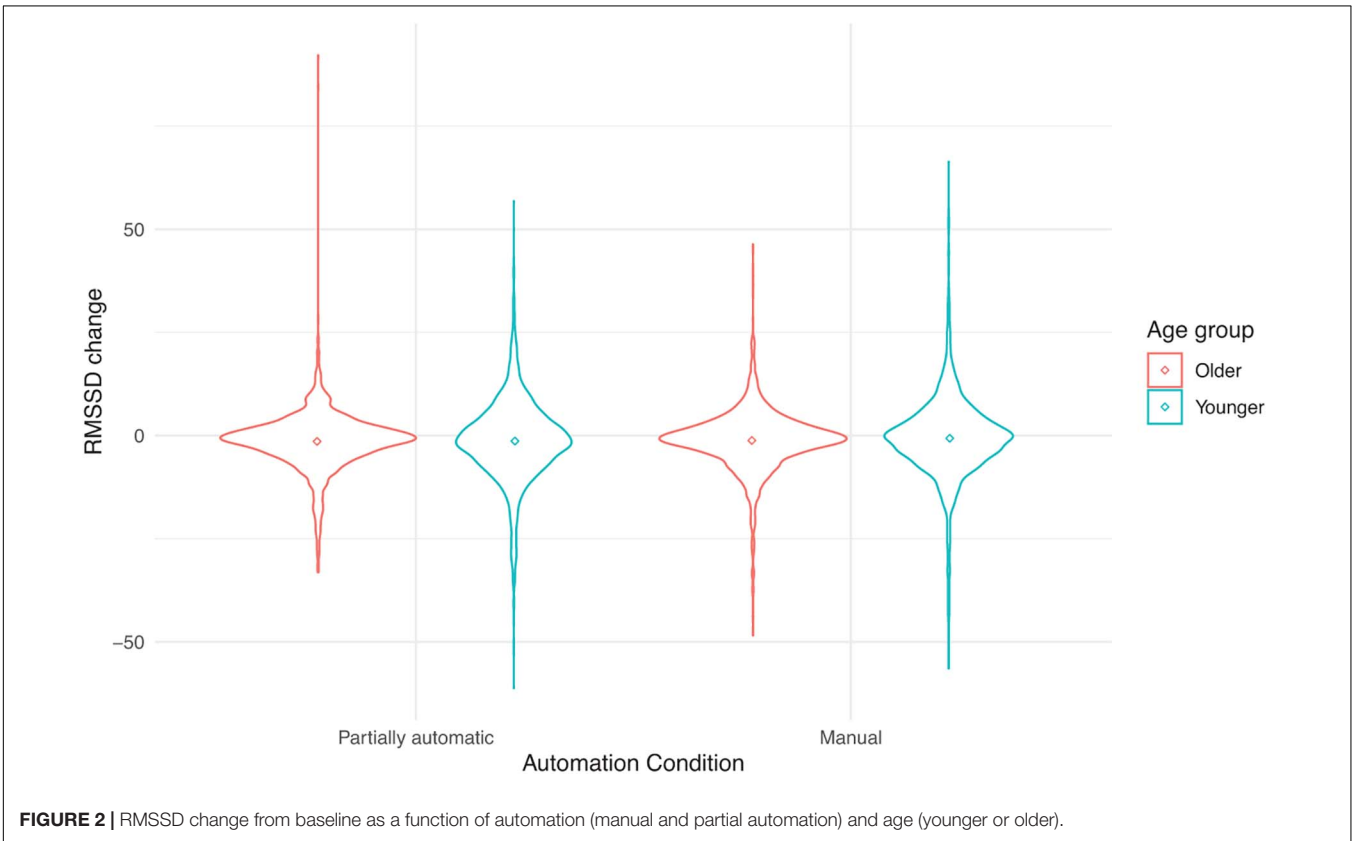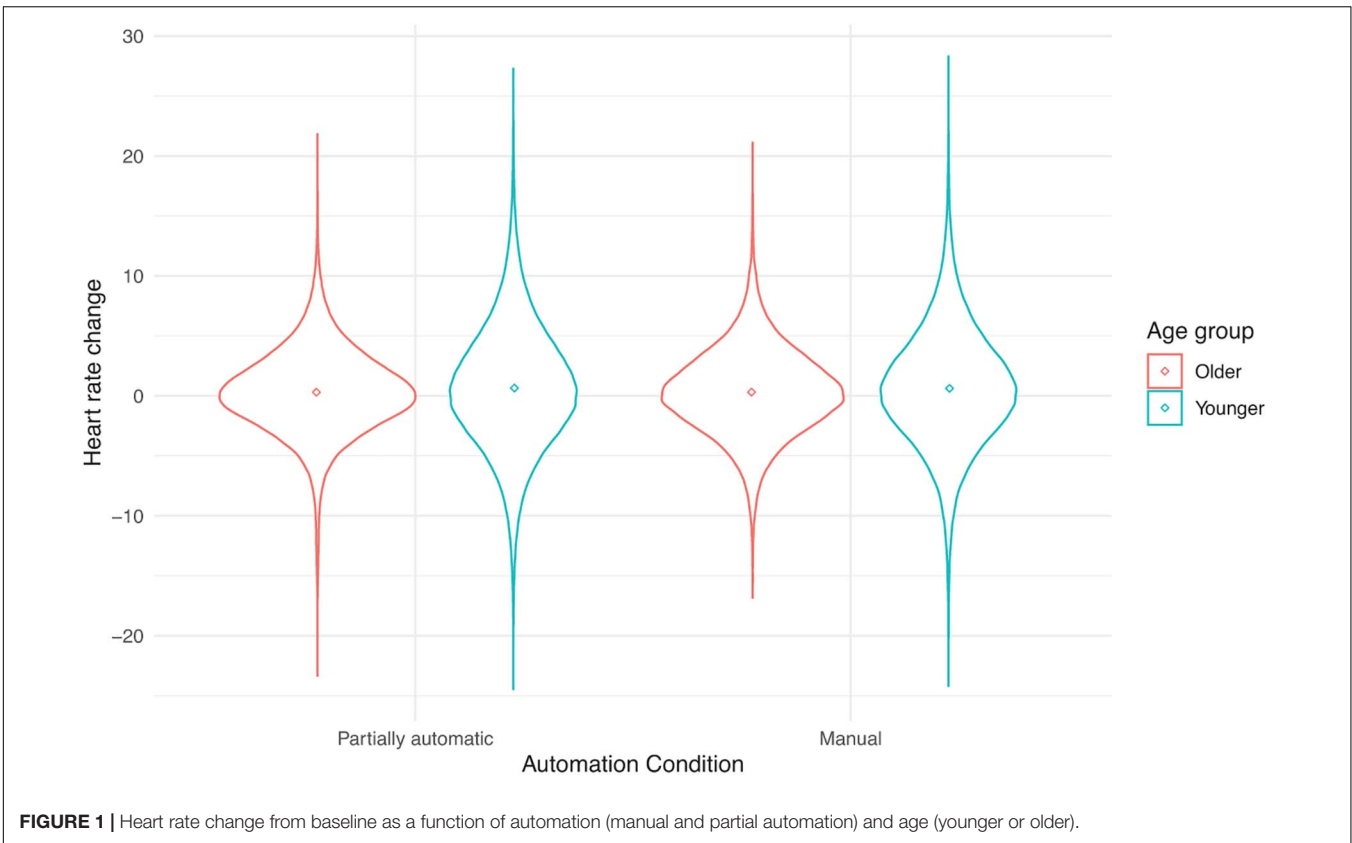### The Effect of Automation on Heart Rate

A linear mixed-effects model was run with a baseline-corrected heart rate as the outcome. Neither the main effect of automation, $F(1, 320.48) = 0.05$, $p = 0.82$, age, $F(1, 60.71) = 0.69$, $p = 0.41$, nor vehicle, $F(3, 355.24) = 0.42$, $p = 0.74$ were significant. Likewise, none of the 2-way or 3-way interactions were significant.
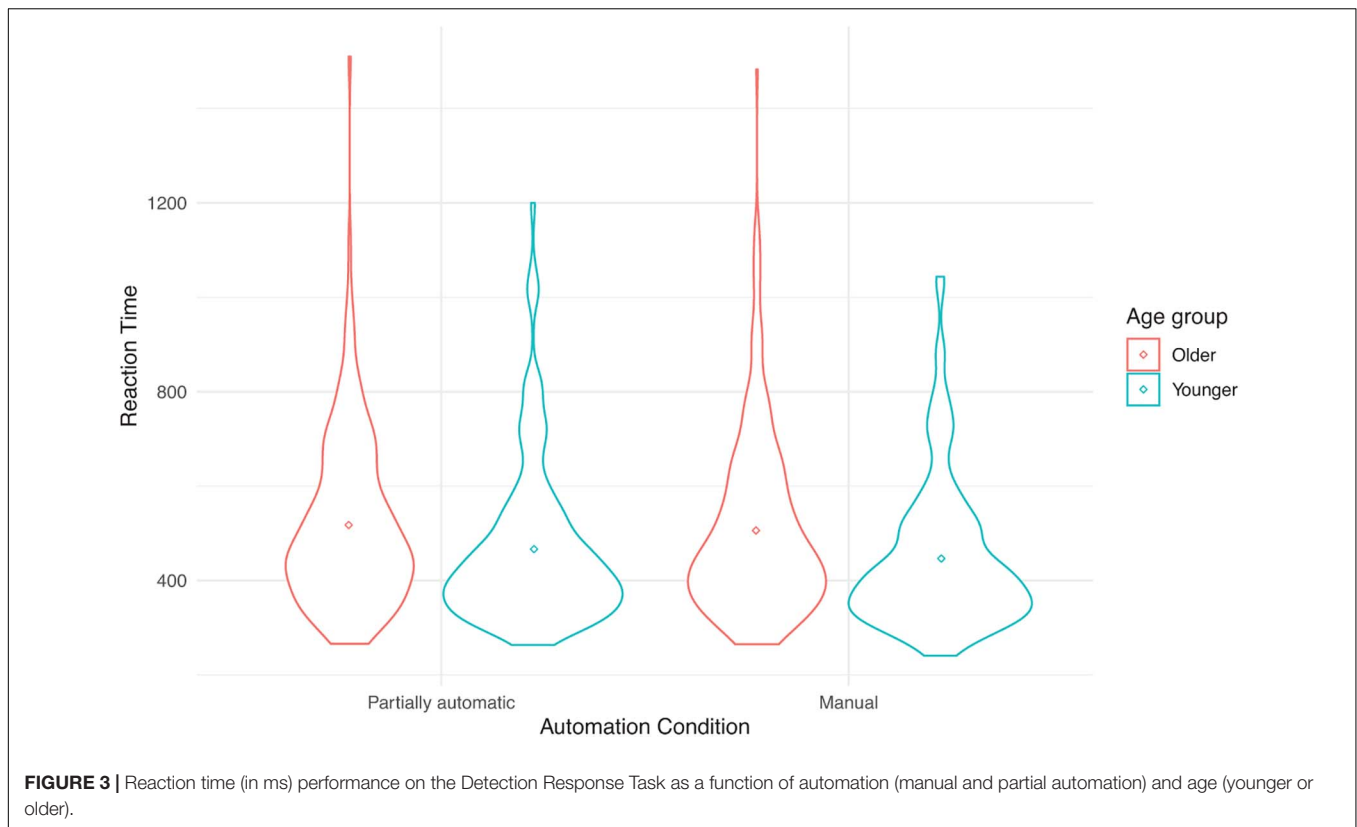
The Effect of Automation on RMSSD

A linear mixed-effects model examined how RMSSD varied as a function of automation (manual and partially automated), age (young or late-middle-aged), and vehicle (Cadillac, Nissan Rogue, Tesla, and Volvo), and participants as the random intercept. No main effect of automation was found, $F(1, 321.99) = 1.76$, $p = 0.19$. In addition, no main effect of age, $F(1, 60.96) = 0.04$, $p = 0.84$, or vehicle was found, $F(3, 358.84) = 1.45$,

TABLE 1 | Means (and Standard Error) for heart rate, RMSSD, and DRT reaction time as a function of automation (manual and partial automation), age (younger or older), and vehicle (Cadillac, Nissan Rogue, Tesla, and Volvo).

Measure	Automation	Young				Old			
		Cadillac	Nissan	Tesla	Volvo	Cadillac	Nissan	Tesla	Volvo
Heart rate	Manual	0.49 (0.32)	0.37 (0.32)	0.76 (0.34)	0.81 (0.33)	0.58 (0.34)	−0.11 (0.34)	0.38 (0.33)	0.60 (0.34)
	Partial automation	0.26 (0.33)	1.05 (0.32)	0.38 (0.33)	0.48 (0.33)	0.30 (0.34)	0.27 (0.33)	0.33 (0.34)	0.55 (0.34)
RMSSD	Manual	−0.45 (0.96)	−0.28 (0.96)	0.07 (1.01)	−2.02 (0.98)	0.68 (1.02)	−0.64 (0.98)	−2.72 (0.96)	−0.96 (1.00)
	Partial automation	−0.49 (0.97)	−1.90 (0.96)	−0.66 (0.99)	−2.36 (0.97)	−1.83 (1.02)	−0.25 (0.98)	−1.81 (1.0)	−1.71 (1.0)
DRT reaction time	Manual	437 (28.5)	443 (28.9)	455 (29.2)	480 (28.9)	498 (31.1)	495 (30.4)	498 (30.5)	523 (31.1)
	Partial automation	459 (28.5)	463 (28.9)	468 (29.2)	489 (28.9)	506 (31.4)	529 (30.4)	508 (30.5)	518 (31.1)

FIGURE 1 | Heart rate change from baseline as a function of automation (manual and partial automation) and age (younger or older).

FIGURE 2 | RMSSD change from baseline as a function of automation (manual and partial automation) and age (younger or older).

FIGURE 3 | Reaction time (in ms) performance on the Detection Response Task as a function of automation (manual and partial automation) and age (younger or older).

$p = 0.23$. Neither the automation by age by vehicle interaction nor any of the 2-way or 3-way interactions were significant.

The Effect of Automation on DRT Reaction Time

A linear mixed-effects model with reaction time as the outcome did not have a significant main effect of automation, $F(1, 325.77) = 2.76$, $p = 0.10$. Similarly, neither the main effects of age, $F(1, 73.77) = 1.77$, $p = 0.19$, nor vehicle, $F(3, 337.02) = 1.67$, $p = 0.17$ were significant. The 2-way and 3-way interactions were also not significant.

Bayes Factor Analyses

In order to examine the effect of automation on RMSSD, heart rate, and DRT reaction time, for each of these outcomes, a Bayes factor analysis was conducted by running a full model with main effects and interactions of automation, age, vehicle, and participants as the random intercept. Next, a restricted model was run without automation with main and interaction effects of age and vehicles as predictors and participants as the random intercept. These full and restricted models were compared to calculate a Bayes Factor of 0.0002, 0.0004, and 0.0002 for heart rate, RMSSD, and DRT reaction time, respectively (see **Figure 4**). According to the Lee and Wagenmakers (2014) classification scheme for interpreting Bayes factors, these values suggest extreme evidence that favors the null hypothesis for the effect of automation on heart rate, RMSSD, and DRT reaction time.
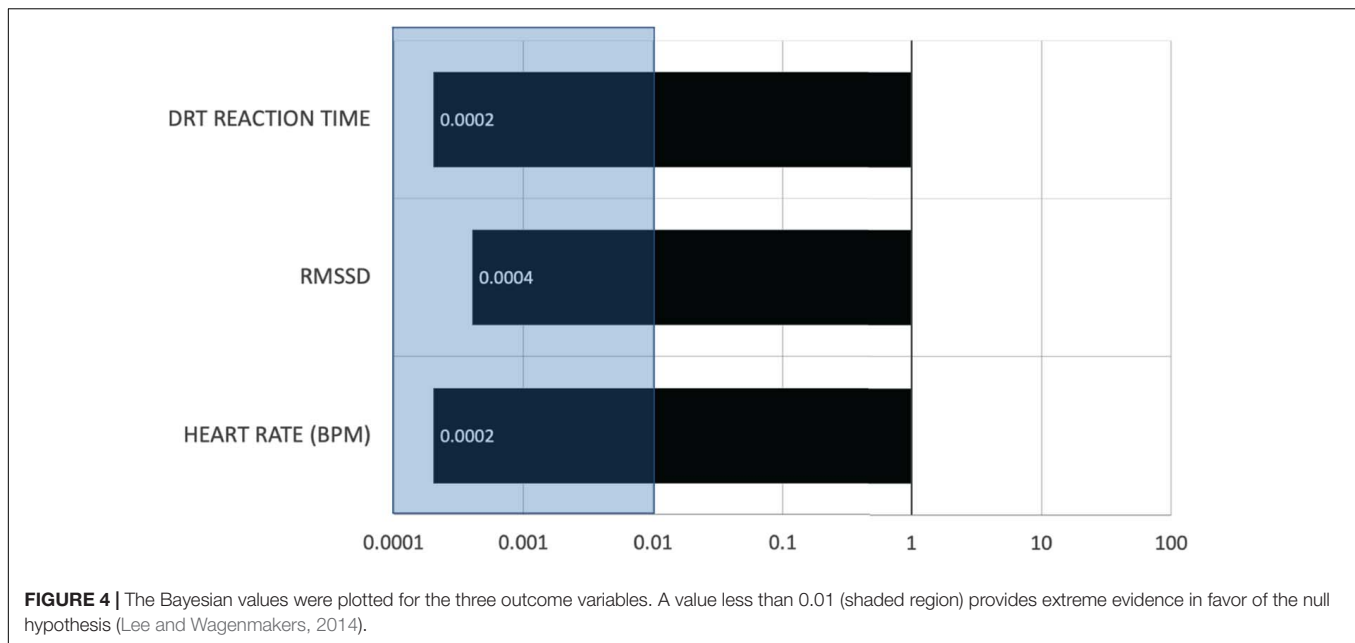
Self-Reported Driving Experience

All participants drove in the partially automated conditions and their experiences on a 5-point scale (completed disagree to completely agree) were analyzed by comparing the responses to the midpoint. The results are reported in **Table 2**. Participants could relax, but relative to manual driving, they were neither less stressed nor bored when the automated driving systems were activated. Participants reported not engaging in unrelated activities while driving the automated systems, such as daydreaming. Participants were well-calibrated in their trust in the automated features of the vehicles. On the one hand, they believed that it made traveling safer and enjoyable, and they were not more anxious or nervous while driving it relative to manual driving. At the same time, they showed restraint when the driving conditions were challenging (e.g., curvy/hilly roads and heavy traffic).

DISCUSSION

Lack of Differences Between Manual and Partially Automated Modes

In order to examine the possible impact of partially automated technology on drivers' arousal and cognitive demand, it is necessary to have sensitive near real-time measures that can detect changes in real-world driving tasks. The current study used heart rate (an arousal measure), RMSSD (a heart rate variability based cognitive demands measure), and DRT (a

FIGURE 4 | The Bayesian values were plotted for the three outcome variables. A value less than 0.01 (shaded region) provides extreme evidence in favor of the null hypothesis (Lee and Wagenmakers, 2014).

TABLE 2 | Results from the driving experiences questionnaire.

Items	Finding	Statistic	Mean (SE)
The driving experiences related questions			
I was able to relax when the automated driving systems were on	Agree	$t(210) = 5.32, p < 0.001$	3.44 (0.08)
I was able to engage in more activities unrelated to driving when the automated driving systems were on	Disagree	$t(210) = 6.35, p < 0.001$	2.43 (0.09)
The automated driving system made traveling boring for me	Disagree	$t(210) = 8.95, p < 0.001$	2.26 (0.08)
The automated driving systems made traveling safer	Agree	$t(210) = 5.46, p < 0.001$	3.39 (0.07)
I was anxious and nervous when the automated driving systems were on	Disagree	$t(210) = 5.69, p < 0.001$	2.51 (0.09)
The automated driving system made traveling more enjoyable	Agree	$t(210) = 6.14, p < 0.001$	3.47 (0.08)
The automated driving system took the fun out of driving	Disagree	$t(210) = 5.59, p < 0.001$	2.49 (0.09)
The automated driving system allowed me to think and daydream	Disagree	$t(210) = 4.53, p < 0.001$	2.61 (0.09)
I was uncomfortable relinquishing control of the vehicle to the automated driving system on curvy and hilly roads	Agree	$t(210) = 5.31, p < 0.001$	3.48 (0.09)
I was uncomfortable relinquishing control of the vehicle to the automated driving systems in heavier traffic	Agree	$t(210) = 3.08, p = 0.002$	3.28 (0.09)
I was concerned that the automated driving systems would shut off unexpectedly	Not sig.	$t(210) = 0.10, p = 0.919$	2.99 (0.09)
The automated driving system reduced the stress of driving	Not sig.	$t(210) = 1.84, p = 0.067$	3.16 (0.07)
Intentions to use and purchase automated driving systems			
I would not feel comfortable using automated driving systems on most roads	Disagree	$t(210) = 7.31, p < 0.001$	2.39 (0.08)
If I was tired or distracted, I would rely heavily on automated driving systems	Disagree	$t(210) = 2.46, p = 0.015$	2.78 (0.09)
I would utilize the automated driving systems in a vehicle as much as possible	Agree	$t(210) = 8.41, p < 0.001$	3.68 (0.08)
I would not feel comfortable using the automated driving systems in a vehicle without monitoring it closely	Agree	$t(210) = 13.16, p < 0.001$	4.00 (0.08)
If I can afford it, I am going to buy or lease a car with automated driving systems	Agree	$t(210) = 7.84, p < 0.001$	3.67 (0.09)
I am going to make sure that the next car I buy, or lease has automated driving systems	Not sig.	$t(210) = 1.84, p = 0.067$	3.15 (0.08)

behavioral performance task based cognitive demands measure) to compare differences in drivers' arousal and cognitive demands during manual versus partially automated driving. To our knowledge, this study is the first effort to examine the partial vehicle automation on arousal and cognitive demands with a representative sample of younger and late-middle-aged drivers and vehicles on real highways while their psychophysiological and behavioral responses were assessed in real-time.

The results suggested that there were no differences in heart rate and its variability or DRT reaction time performance between manual and partially automated modes of driving either in younger or late-middle-aged adults across the four test vehicles. A Bayes Factor analysis on heart rate, RMSSD, and DRT reaction time data showed *extreme evidence in favor of the null hypothesis,* suggesting that drivers new to partial automation maintain comparable levels of arousal and cognitive demands during manual and partially automated driving as no evidence of under or over-arousal and cognitive demands was found. In another forthcoming paper (McDonnell et al., in review), EEG alpha power and frontal theta (a central physiology-based

index of visual engagement and mental workload) were also found to show no significant difference between manual and partially automated modes. Past research on potential cognitive differences between manual and partial automation has been mixed. Various methodological limitations can explain these mixed findings, such as limited sample size, use of only self-report measures, and the use of driving simulators or a single-vehicle. The current study allowed for a fair test of the null hypothesis by testing a large representative sample of drivers and vehicles with sufficient sample size and a combination of reliable physiological and behavioral measures.

The current findings support and extend previous research that has found no difference between manual and partial automation (e.g., Sibi et al., 2017; Stapel et al., 2019; Calvi et al., 2020; Lohani et al., 2020; Várhelyi et al., 2020). Moreover, the pilot study that examined young drivers operating three vehicles on the highway also showed a similar pattern (Lohani et al., 2020). Similar to the current study, the pilot study found no differences across outcomes (heart rate, RMSSD, EEG alpha and theta power, and DRT performance). Bayes Factor analysis had suggested that there was strong evidence that arousal and cognitive demands did not differ during manual and partially automated driving (Lohani et al., 2020). With a single measure, null results are harder to interpret, but replicable effects with multiple reliable outcomes provide a more convincing interpretation of the null hypothesis. Taken together, these findings provide strong evidence that arousal and cognitive demands are similar during manual versus partially automated conditions for drivers who are new to partially automated technology.

A lack of differences between manual and partially automated modes also suggests that the cognitive demands imposed during manual driving are comparable to those imposed by monitoring partial automation. Based on the Neurovisceral Integration model (Thayer et al., 2009), these findings would imply that neural activity associated with cognitive functioning during manual and partially automated driving is comparable. Self-reports from participants suggested that participants were cognitively engaged in the driving process. This may be because participants still have to monitor ongoing traffic conditions and maintain cognitive readiness to take control of the vehicle. Interestingly, the age group did not lead to any differences in heart rate and RMSSD (after accounting for pre-condition baseline) or DRT reaction time. This implies that late-middle-aged drivers are able to use partially automated technology similar to their younger counterparts and that they can benefit from assistance provided by partially automated technology without additional cognitive costs. While this study investigated drivers between 21 and 64 years of age, future research should examine potential differences in teenage drivers and drivers older than 64 years of age to better understand driving partial automation vehicles across the lifespan.

Limitations and Outstanding Questions

There are a few limitations of the current study. First, this study focused on drivers with no prior experience with partial driving automation. Drivers that are more experienced with partially automated technology may respond differently to automated vehicles than those with limited prior experience. The perceived workload was reduced for automation-experienced drivers, while it did not change from the manual mode for inexperienced drivers (Stapel et al., 2019). Motorists may learn to accept and trust the automation technology with additional driving experience (Beggiato et al., 2015). Experienced drivers may get better at calibrating their trust after understanding the automation system's limitations (Walker et al., 2018). However, it is also possible that more trusting participants increase their reliance on automation (Walker et al., 2019), resulting in poor readiness to switch from automated to manual driving safely. More work is needed to explore any associated risks.

Second, we had a research assistant sit in the passenger seat next to the participant for safety reasons, and we cannot rule out that this presence may have impacted drivers' performance and physiology. However, any confounding effects of social presence were constant in both manual and partial automation conditions across all the participants. Third, this study was not explicitly designed to test gender differences. Factors such as the different phases of the menstrual cycle were not recorded or accounted for and could affect the RMSSD data. Future work is needed to understand the effect this may have on RMSSD. Fourth, this study did not examine traffic conditions, and low and high traffic demands could moderate the outcomes. Finally, it is possible that constantly transitioning between manual and partially automated driving modes could be a demanding task for some motorists. For instance, one study found that older adults were slower at switching between manual and partially automated modes (Wu et al., 2019). Future research should examine transition-related driving demands on motorists while driving partially automated vehicles.

Benefits and Challenges of Adopting HRV Indices in Driving Safety

With advances in methodological developments, it is now possible to collect high-quality psychophysiological data outside traditional lab settings from research-grade equipment at a low cost. This study conducted on real roads highlights the applicability of heart rate variability to real-world automation driving research. However, some caution is warranted while interpreting HRV measurement. As is true for most psychophysiological measures, HRV does not have a one-to-one correspondence with a single psychological construct (e.g., Cacioppo and Tassinary, 1990). HRV may be sensitive to many cognitive factors that may occur in real-world driving contexts. For example, over time, drivers may become more relaxed or disengaged with the driving task and both scenarios would lead to a lower-arousal and a corresponding increase in HRV indices (Jasper et al., 2016). Furthermore, there may be other factors (e.g., Laborde et al., 2017) while driving that can influence HRV measurements, such as driving task-related factors (e.g., bad traffic), driver-related factors (e.g., a driver has irregular heartbeats or is on psychotropic medications), or concurrent activities (talking, smoking, or caffeine intake). Because such

factors may influence HRV indices and may vary across driving conditions; thoughtful analyses of such effects in conjunction with other measures (such as behavioral performance) is crucial for accurate interpretations.

A challenge with HRV measurement is that it is susceptible to artifacts that can result in inaccurate values. Some analytic approaches have adopted shorter segments (e.g., 30 s) of data to detect momentary effects of driving-related workload (Stuiver et al., 2014). While preliminary solutions to handle noise and artifacts have been proposed (e.g., Nowara et al., 2018; van Gent et al., 2019; Brown et al., 2020), more work is needed to develop near-real-time artifact detection methods that are amenable to driving task-related changes. As in the current study, we argue that a multi-method approach provides a more accurate interpretation of cognitive demands in applied settings. Several efforts in driving research have proposals to assess cognitive states using a multi-method approach that seems promising (Fu et al., 2016; Brouwer et al., 2017; Aricò et al., 2018; Haouij et al., 2018; Paredes et al., 2018; Rastgoo et al., 2018). The current findings imply that RMSSD is a suitable measure that could be utilized in multi-method studies to evaluate drivers' cognitive demand.

CONCLUSION

This investigation was conducted on real roads with a large and representative sample of younger and late-middle-aged motorists to compare manual and partially automated driving. The current findings show important evidence of no difference in arousal and cognitive demands (extreme evidence in favor of the null hypothesis), suggesting that drivers maintain similar levels of arousal and cognitive demands during manual and partially automated driving. These theoretically relevant results imply that younger and late-middle-aged motorists who are new to partial automation may maintain arousal and cognitive levels comparable to manual driving. This research extends a growing literature highlighting the applicability of heart rate variability to real-world automation research.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board, University of Utah. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DS and JC conceived the study idea and design. ML, GE, TS, AC, and KC were involved in data collection and processing. ML analyzed the data and wrote the first draft of the manuscript. DS, JC, and AM provided suggestions to improve the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., and Babiloni, F. (2018). Passive BCI beyond the lab: current trends and future directions. *Physiol. Meas.* 39:08TR02. doi: 10.1088/1361-6579/aad57e

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Beggiato, M., Pereira, M., Petzoldt, T., and Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transp. Res. Part F* 35, 75–84. doi: 10.1016/j.trf.2015.10.005

Bengler, K., Kohlmann, M., and Lange, C. (2012). Assessment of cognitive workload of in-vehicle systems using a visual peripheral and tactile detection task setting. *Work* 41(Suppl. 1), 4919–4923. doi: 10.3233/WOR-2012-0786-4919

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34, 623–648. doi: 10.1111/j.1469-8986.1997.tb02140.x

Berntson, G. G., Quigley, K. S., and Lozano, D. (2007). Cardiovascular psychophysiology. *Handb. Psychophysiol.* 3, 182–210. doi: 10.1017/CBO9780511546396.008

Biondi, F. N., Lohani, M., Hopman, R., Mills, S., Cooper, J. M., and Strayer, D. L. (2018). "80 MPH and out-of-the-loop: effects of real-world partially automated driving on driver workload and arousal," in *Proceedings of the Human Factors*

and Ergonomics Society Annual Meeting, Vol. 62, (Los Angeles, CA: SAGE Publications), 1878–1882. doi: 10.1177/1541931218621427

Broadhurst, P. L. (1959). The interaction of task difficulty and motivation: the Yerkes Dodson Law revived. *Acta Psychol.* 16, 321–338. doi: 10.1016/0001-6918(59)90105-2

Brouwer, A., Snelting, A., Jaswa, M., Flascher, O., Krol, L., and Zander, T. (2017). "Physiological effects of adaptive cruise control behaviour in real driving," in *Proceedings of the 2017 ACM Workshop on An Application-oriented Approach to BCI out of the Laboratory*, Limassol. doi: 10.1145/3038439.3038441

Brown, S. B. R. E, Brosschot, J. F., Versluis, A., Thayer, J. F., and Verkuil, B. (2020). Assessing new methods to optimally detect episodes of non-metabolic heart rate variability reduction as an indicator of psychological stress in everyday life: a thorough evaluation of six methods. *Front. Neurosci.* 14:564123. doi: 10.3389/fnins.2020.564123

Bruyas, M. P., and Dumont, L. (2013). "Sensitivity of Detection Response Task (DRT) to the driving demand and task difficulty," in *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. Symposium presented at Bolton Landing*, New York, NY, 17–20.

Cacioppo, J. T., and Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *Am. Psychol.* 45, 16–28. doi: 10.1037/0003-066X.45.1.16

Calvi, A., D'Amico, F., Ciampoli, L. B., and Ferrante, C. (2020). Evaluation of driving performance after a transition from automated to manual control: a

driving simulator study. *Transp. Res. Procedia* 45, 755–762. doi: 10.1016/j.trpro.2020.02.101

Cisler, D., Greenwood, P. M., Roberts, D. M., McKendrick, R., and Baldwin, C. L. (2019). Comparing the relative strengths of EEG and low-cost physiological devices in modeling attention allocation in semiautonomous vehicles. *Front. Hum. Neurosci.* 13:109. doi: 10.3389/fnhum.2019.00109

Cooper, J. M., Castro, S. C., and Strayer, D. L. (2016). "Extending the Detection Response Task to simultaneously measure cognitive and visual task demands," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 60, Prague, 1962–1966. doi: 10.1177/1541931213601447

Fu, R., Wang, H., and Zhao, W. (2016). Dynamic driver fatigue detection using hidden Markov model in real driving condition. *Expert Syst. Appl.* 63, 397–411. doi: 10.1016/j.eswa.2016.06.042

Gershon, P., Shinar, D., and Ronen, A. (2009). Evaluation of experience-based fatigue countermeasures. *Accid. Anal. Prev.* 41, 969–975. doi: 10.1016/j.aap.2009.05.012

Haouij, N. E., Poggi, J.-M., Sevestre-Ghalila, S., Ghozi, R., and Jaïdane, M. (2018). "AffectiveROAD system and database to assess driver's attention," in *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, Pau. doi: 10.1145/3167132.3167395

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychol. Rev.* 62, 243–254. doi: 10.1037/h0041823

Heikoop, D. D., de Winter, J. C., van Arem, B., and Stanton, N. A. (2019). Acclimatizing to automation: driver workload and stress during partially automated car following in real traffic. *Transp. Res. Part F* 65, 503–517. doi: 10.1016/j.trf.2019.07.024

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

International Organization for Standardization [ISO] (2016). *Road Vehicles – Transport Information and Control Systems – Detection Response Task (DRT) for Assessing Attentional Effects of Cognitive Load in Driving (Rep. ISO 17488)*. Geneva: International Organization for Standardization.

Jasper, P., Sibley, C., and Coyne, J. (2016). "Using heart rate variability to assess operator mental workload in a command and control simulation of multiple unmanned aerial vehicles," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 60, (Los Angeles, CA: SAGE Publications), 1125–1129. doi: 10.1177/1541931213601264

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kim, J., Revell, K., Langdon, P., Bradley, M., Politis, I., Thompson, S., et al. (2020). "Drivers' interaction with, and perception toward semi-autonomous vehicles in naturalistic settings," in *Proceedings of the International Conference on Intelligent Human Systems Integration Springer*, Cham, 20–26. doi: 10.1007/978-3-030-39512-4_4

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspect. Psychol. Sci.* 6, 299–312. doi: 10.1177/1745691611406925

Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting. *Front. Psychol.* 8:213. doi: 10.3389/fpsyg.2017.00213

Lee, H. B., Kim, J. S., Kim, Y. S., Baek, H. J., Ryu, M. S., and Park, K. S. (2007). "The relationship between HRV parameters and stressful driving situation in the real road," in *Proceedings of the 6th International Special Topic Conference on Information Technology Applications in Biomedicine, 2007*, Tokyo. doi: 10.1109/ITAB.2007.4407380

Lee, M. D., and Wagenmakers, E. J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge: Cambridge University Press.

Lenneman, J. K., and Backs, R. W. (2009). Cardiac autonomic control during simulated driving with a concurrent verbal working memory task. *Hum. Fact.* 51, 404–418. doi: 10.1177/0018720809337716

Liang, W. C., Yuan, J., Sun, D. C., and Lin, M. H. (2009). Changes in physiological parameters induced by indoor simulated driving: effect of lower body exercise at mid-term break. *Sensors* 9, 6913–6933. doi: 10.3390/s90906913

Lohani, M., Cooper, J. M., Erickson, G., Simmons, T., McDonnell, A., Crabtree, K., et al. (2020). *Driver Arousal and Workload Under Partial Vehicle Automation:*

A Pilot Study. Proceedings of the Human Factors Ergonomic Society. Available online at: https://journals.sagepub.com/doi/abs/10.1177/1071181320641471?journalCode=proe (accessed June 1, 2020).

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13:57. doi: 10.3389/fnhum.2019.00057

Malik, M. (1996). Heart rate variability. 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. *Eur. Heart J.* 17, 354–381.

Mauss, I. B., and Robinson, M. D. (2010). Measures of emotion: a review. *Cogn. Emot.* 23, 209–237, doi: 10.1080/02699930802204677

McDonnell, A. S., Simmons, T. G., Erickson, G. G., Lohani, M., Cooper, J. M., and Strayer, D. L. (in review). This is your brain on autopilot: neural indices of driver workload and engagement during partial vehicle automation.

Mehler, B., Reimer, B., and Wang, Y. (2011). "A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload," in *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design: Driving Assessment, 2011*, Lake Tahoe, CA, 590–597. doi: 10.17077/drivingassessment.1451

Morey, R. D., Rouder, J. N., and Jamil, T. (2018). *BayesFactor: Computation of Bayes Factors for common designs. R package version 0.9. 12-4.2. Computer software*. Available online at: https://CRAN.R-project.org/package=BayesFactor (accessed June 1, 2020).

Mulder, G. (1985). "Attention, effort and sinus arrhythmia: how far are we," in *Psychophysiology of Cardiovascular Control*, eds J. F. Orlebeke, G. Mulder, and L. P. J. van Doornen (New York, NY: Plenum Press), 407–424.

Mulder, G., and Mulder, L. J. (1981). Information processing and cardiovascular control. *Psychophysiology* 18, 392–402. doi: 10.1111/j.1469-8986.1981.tb02470.x

Noda, A., Miyaji, M., Wakuda, Y., Hara, Y., and Yasuma, F. (2015). Simultaneous measurement of heart rate variability and blinking duration to predict sleep onset and drowsiness in drivers. *J. Sleep Disord. Ther.* 4, 2167–2277. doi: 10.4172/2167-0277.1000213

Nowara, E. M., Marks, T. K., Mansour, H., and Veeraraghavan, A. (2018). "SparsePPG: towards driver monitoring using camera-based vital signs estimation in near-infrared," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT. doi: 10.1109/CVPRW.2018.00174

Paredes, P. E., Ordonez, F., Ju, W., and Landay, J. A. (2018). "Fast & furious: detecting stress with a car steering wheel," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal, QC. doi: 10.1145/3173574.3174239

Peltola, M. A. (2012). Role of editing of R-R intervals in the analysis of heart rate variability. *Front. Physiol.* 23:148. doi: 10.3389/fphys.2012.00148

Piotrowski, Z., and Szypulska, M. (2017). Classification of falling asleep states using HRV analysis. *Biocybern. Biomed. Eng.* 37, 290–301. doi: 10.1016/j.bbe.2017.02.003

Quintana, D. S., and Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. *BMC Psychiatry* 18:178. doi: 10.1186/s12888-018-1761-4

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/.

Rastgoo, M. N., Nakisa, B., Rakotonirainy, A., Chandran, V., and Tjondronegoro, D. (2018). A critical review of proactive detection of driver stress levels based on multimodal measurements. *ACM Comput. Surv.* 51, 1–35. doi: 10.1145/3186585

Robbins, T. W., and Everitt, B. J. (1995). "Arousal systems and attention," in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, MA: The MIT Press), 703–720.

SAE (2016). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles (Surface Vehicle Recommended Practice: Superseding J3016 Jan 2014), SAE International*. Available online at: https://www.sae.org/standards/content/j3016_201806/. (accessed June 1, 2020).

Satpute, A. B., Kragel, P. A., Barrett, L. F., Wager, T. D., and Bianciardi, M. (2019). Deconstructing arousal into wakeful, autonomic and affective varieties. *Neurosci. Lett.* 693, 19–28. doi: 10.1016/j.neulet.2018. 01.042

Shaffer, F., McCraty, R., and Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Front. Psychol.* 5:1040. doi: 10.3389/fpsyg.2014.01040

Sibi, S., Baiters, S., Mok, B., Steiner, M., and Ju, W. (2017). "Assessing driver cortical activity under varying levels of automation with functional near infrared spectroscopy," in *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, 1509–1516.

Smith, R., Thayer, J. F., Khalsa, S. S., and Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. *Neurosci. Biobehav. Rev.* 75, 274–296. doi: 10.1016/j.neubiorev.2017.02.003

Solís-Marcos, I., Ahlström, C., and Kircher, K. (2018). Performance of an additional task during Level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation. *Hum. Factors* 60, 778–792. doi: 10.1177/0018720818773636

Stapel, J., Mullakkal-Babu, F. A., and Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transp. Res. Part F* 60, 590–605. doi: 10.1016/j.trf.2018. 11.006

Stuiver, A., Brookhuis, K. A., de Waard, D., and Mulder, B. (2014). Short- term cardiovascular measures for driver support: increasing sensitivity for detecting changes in mental workload. *Int. J. Psychophysiol.* 92, 35–41. doi: 10.1016/j. ijpsycho.2014.01.010

Sugie, R., Arakawa, T., and Kozuka, K. (2016). Detection of fatigue in long-distance driving by heart rate variability. *Innovat. Comput. Inform. Control Exp. Lett.* 10, 1553–1559.

Thayer, J. F., Hansen, A. L., Saus-Rose, E., and Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Ann. Behav. Med.* 37, 141–153. doi: 10.1007/s12160-009-9101-z

van Gent, P., Farah, H., van Nes, N., and van Arem, B. (2019). Analysing noisy driver physiology real-time using off-the-shelf sensors: heart rate analysis software from the taking the fast lane project. *J. Open Res. Softw.* 7:32. doi: http://doi.org/10.5334/jors.241

Várhelyi, A., Kaufmann, C., Johnsson, C., and Almqvist, S. (2020). Driving with and without automation on the motorway–an observational study. *J. Intell. Transp. Syst.* 1–22.

Walker, F., Boelhouwer, A., Alkim, T., Verwey, W. B., and Martens, M. H. (2018). Changes in trust after driving level 2 automated cars. *J. Adv. Transp.* 2018:1045186. doi: 10.1155/2018/1045186

Walker, F., Wang, J., Martens, M. H., and Verwey, W. B. (2019). Gaze behaviour and electrodermal activity: objective measures of drivers' trust in automated vehicles. *Transp. Res. Part F* 64, 401–412.

Wekselblatt, J. B., and Niell, C. M. (2015). Behavioral state—getting in the zone. *Neuron* 87, 7–9. doi: 10.1016/j.neuron.2015.06.020

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* 3, 159–177.

Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., and Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: a simulator study towards understanding drivers in automated driving. *Accid. Anal. Prev.* 124, 202–209.

Yerkes, R. M., and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.* 18, 459–482. doi: 10.1002/cne.920180503

Young, R. A., Hsieh, L., and Seaman, S. (2013). "The tactile detection response task: preliminary validation for measuring the attentional effects of cognitive load," in *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, New York, NY, 71–77. doi: 10.17077/drivingassessment.1469

Zhai, J., and Lu, G. (2019). "A study of how drivers' subjective workload and driving performance change under varying levels of automation and critical situations," in *Proceedings of the 19th COTA International Conference of Transportation Professionals*, Nanjing, 344–355.