Acclimatizing to automation: driver workload and stress during partially automated car following in real traffic

Daniël D. Heikoop[a1], Joost C. F. de Winter[b], Bart van Arem[c], and Neville A. Stanton[a]

[1]Corresponding author
[a]Transportation Research Group, Faculty of Engineering and the Environment, Boldrewood Innovation Campus, University of Southampton, Burgess Road, Southampton, SO16 7QF, United Kingdom
[b]Department of BioMechanical Engineering, Delft University of Technology, the Netherlands
[c]Department of Transport & Planning, Delft University of Technology, the Netherlands

Automated driving systems are increasingly prevalent on public roads, but there is currently little knowledge on the level of workload and stress of drivers operating an automated vehicle in a real environment. The present study aimed to measure driver workload and stress during partially automated driving in real traffic. We recorded heart rate, heart rate variability, respiratory rate, and subjective responses of nine test drivers in the Tesla Model S with Autopilot. The participants, who were experienced with driver assistance systems but naïve to the Tesla, drove a 32 min motorway route back and forth while following a lead car in regular traffic. In one of the two drives, participants performed a heads-up detection task of bridges they went underneath. Averaged across the two drives, the participants' mean self-reported overall workload score on the NASA Task Load Index was 19%. Moreover, the participants showed a reduction of heart rate and self-reported workload over time, suggesting that the participants became accustomed to the experiment and technology. The mean hit (i.e., pressing the button near a bridge) rate in the detection task was 88%. In conclusion, driving with the Tesla Autopilot on a motorway involved a low level of workload that decreased with time on task.

1. Introduction

1.1. Workload and Stress in Automated Driving

Cars that provide combined longitudinal and lateral automated control support have recently been introduced on the market. Automated driving may be expected to reduce workload and stress as compared to manual driving because the driver does not have to control the vehicle. However, unless the driving task is fully automated (SAE level 5), automated driving may cause high workload and stress, because the driver needs to supervise both the human-machine interface and the state of the car in relation to the outside environment (for an illustration, see Fig. 1). More specifically, the driver of an automated car has to remain attentive to reclaim manual control if required [1, 2] a task that may be demanding and stressful [3]. Furthermore, the type of supervisory control shown in Figure 1 may cause out-of-the-loop problems, such as loss of situation awareness and mode errors, which resemble those observed in aviation and process control [4, 5, 6; see also 7]. A survey by Dikmen and Burns [8] among 121 Tesla owners found that automation failures (e.g., failure to detect lanes) were frequent but not perceived as risky. Furthermore, the majority of respondents indicated that it is important to remain alert and to be aware of the automation's limitations.

1.2. Prior research on workload and stress in automated driving.

The majority of Human Factors research on driver workload in automated vehicles has been conducted in driving simulators [see 9 for a review]. Overall, the results indicate that the self-reported workload as assessed with the NASA Task Load Index (TLX) is substantially lower in automated driving than in manual driving [see 9 for a review], and below 50% on a scale from 0 to 100% (see Table 1 for an overview).

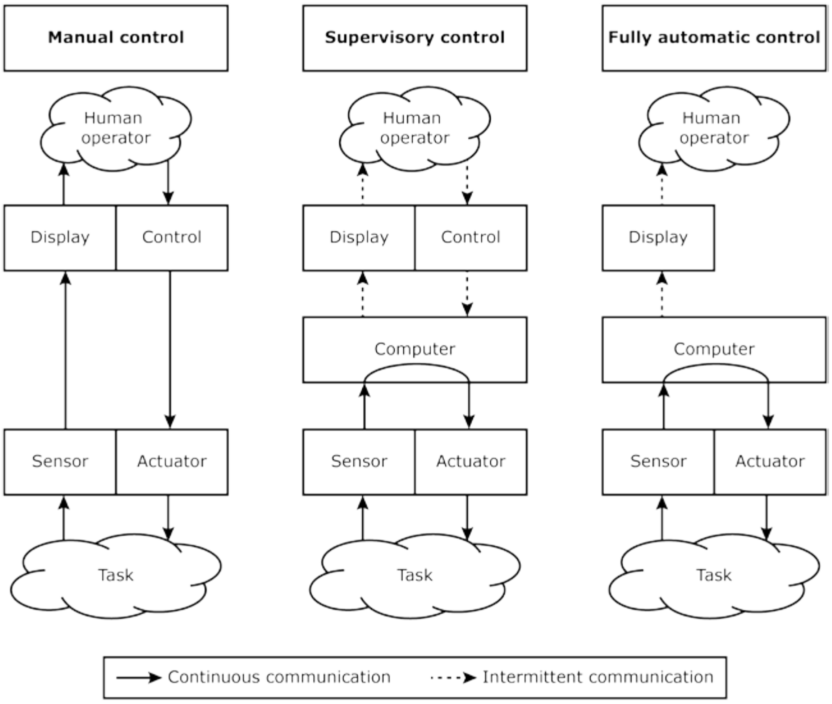


Figure 1. Manual control (left), supervisory control (middle), and fully automatic control (right). In manual control, the driver controls the car via manipulators (steering wheel & pedals) and continuously receives information from the car in the environment (i.e., task). In fully automatic driving, the human has no contribution to the driving task other than to set a destination (or to press an emergency stop button). Hence, the driver is out of the control loop completely. In supervisory control, the driver interacts with a computer that closes the control loop via sensors and actuators, while the driver intermittently (1) provides instructions to the computer, (2) receives information via displays, and (3) receives information from the car in the environment [from 12].

A small number of on-road studies are available. Recently, Endsley [10] conducted a single-subject naturalistic driving study using a Tesla Model S over a six-month period. She reported that her situation awareness increased when using automation, because less focus was needed on controlling the vehicle, and more attention could be devoted to looking at traffic and road signage. However, Endsley also experienced various issues of mode confusion and unexpected automation transitions, as well as loss of attention. Endsley further found that ratings of satisfaction, usefulness, and trust gradually increased from months 1–2 towards months 5–6, which is in line with the results of a longitudinal naturalistic driving study on adaptive cruise control (ACC) with 15 participants [11]. Additionally, overall self-reported workload was low, averaging at about 1.3 during months 1–2 and 1.0 during months 3–4, on a scale from 0 to 5 [10].

Eriksson et al. [13] let 12 test drivers use the Tesla autopilot for about 20 minutes per participant. Participants each experienced approximately 12 automation-to-manual control transitions, and completed the NASA-TLX after the ride. The mean overall workload was 19%. Stapel et al. [14] conducted an on-road highway driving study in which 15 participants used the Tesla Autopilot for about 20 minutes. The authors found overall low levels of workload among participants (between 10% and 43%), with the type of road (busy city ring versus relatively empty highway) and prior experience with the Tesla Model S being moderator variables (Table 1). In another on-road study, Banks and Stanton [15] tested a prototype version of automated longitudinal and lateral control in addition to a driver-initiated auto-overtaking system. These authors found relatively high workload on the NASA-TLX (median of 42%) during 9 minutes of automated driving per participant.

The discrepancy between the results of Banks and Stanton [15] and the findings of Eriksson et al. [13] and Stapel et al. [14] may be caused by the fact that the prototype system tested by Banks and Stanton, which included a heads-up display and offered overtake suggestions, was difficult to use or that participants were still learning how to use it. Because the participants in Banks and Stanton [15] drove only 9 minutes with the automation system, the high workload levels "may be a simple reflection of the fact that these ratings were collected during first-time use of the automated system", p. 393.

McDowell et al. [16] and Davis et al. [17] performed on-road trials with automated military convoys. In these studies, where there was no other traffic and, because they were military experiments, object detection was of primary importance. The results showed that automated driving reduced workload and improved performance in object detection in comparison to manual driving.

On-road studies may be expected to yield higher workload than simulator studies, because the latter involve no physical risks of accidents. However, in some cases, on-road studies actually yielded lower workload than simulator-based studies. For example, the reported workload in Eriksson et al. [13] was 19%, compared to the 33% in the simulator study of Manawadu et al. [18]. This difference might be attributable to the participant pools, added events, or secondary tasks. Specifically, the study of Eriksson et al. [13] involved experienced test drivers and did not include a secondary task; participants were merely required to take over and relinquish control of the vehicle throughout the experiment. In Manawadu et al. [18], critical events were triggered, to which the participants had to respond.

Table 1
Overview of workload measurements presenting a NASA-TLX overall workload (TLX OW) score in automated driving studies.

Reference	Simulator/road	Sample size	Mean TLX OW
Banks and Stanton [15]	On-road	32	42% (median)
Borojeni et al. [19]	Simulator	21	30%
Damböck et al. [20]	Simulator	24	33%
De Winter et al. [21]	Simulator	24	31% (exp. 1)
	Simulator	27	31% (exp. 2)
Eriksson et al. [13]	On-road (Tesla)	12	19%
Eriksson and Stanton [22]	Simulator	26	21%
Heikoop et al. [23]	Simulator	22	28%
Large et al. [24]	Simulator	30	36% (partial automation), 21% (high automation)
Manawadu et al. [18]	Simulator	6 (novices)	36%
	Simulator	6 (experienced)	30%
McDowell et al. [16]	On-road (military)	11	40%
Petermeijer et al. [25]	Simulator	24	28% (with auditory and vibrotactile feedback)
Petermeijer et al. [26]	Simulator	18	22%, 36% (with N-Back task)
Saxby et al. [27]	Simulator	36	34% (exp. 1)
	Simulator	56	27% (exp. 2)
Schwalk et al. [28]	Simulator	24	21%
Stapel et al. [14]	On-road (Tesla)	8 (no experience with Tesla)	25% (empty highway), 43% (city ring)
	On-road (Tesla)	7 (experienced with Tesla)	10% (empty highway), 24% (city ring)
Young [29]	Simulator	18	23%
Young and Stanton [30]	Simulator	12	12% (exp. 1)
	Simulator	12	12% (exp. 2)
Young and Stanton [31]	Simulator	24 (novice drivers)	11%
	Simulator	30 (learner drivers)	13%
	Simulator	30 (expert drivers)	20%
	Simulator	30 (advanced drivers)	24%

1.3. Aim of the present study
The present study aimed to assess whether on-road automated driving with the Tesla Model S alleviates driver workload *over time*. Both the on-road studies of Eriksson et al. [13] and Stapel et al. [14] consisted of approximately 20 min of highway driving with the automation engaged (excluding a familiarization drive) and did not report on temporal effects. Our study consisted of 64 minutes of automated highway driving per participant.

Additionally, in our study, a simple detection task was used to add task demands on top of the regular monitoring demands during automated driving. More specifically, participants were instructed to press a handheld button when driving underneath a bridge. The task of detecting bridges is practically convenient in an on-road experiment because bridges are irregularly spaced, and the locations of bridges are retrievable from Google Maps. This detection task is conceptually similar to the approach taken in a previous platooning experiment in a driving simulator [23]. In Heikoop et al. [23], it was found that the detection task (i.e., to detect red cars on the road) increased self-reported mental demands compared to not performing a detection task. We expected to find a similar effect in this study.

2. Methods

2.1. Participants

Nine participants (seven males, two females) aged between 25 and 47 years (*M* = 35.44; *SD* = 8.26) with 6 to 30 years of self-reported driving experience (*M* = 17.56; *SD* = 8.46) took part in this experiment. The participants were employees of a large automotive company. Eight participants indicated that they drove every day and one participant indicated driving 4–6 days a week. Two participants indicated they drove up to 10,000 miles, five up to 20,000, one up to 30,000, and one up to 50,000 miles in the past year. All participants had completed level-2 driver training, an extended driver training specifically designed for people who drive as part of their job and which serves as a legal requirement for insurance purposes. All participants had driven various supercars before, and had experience with advanced driver assistance systems (e.g., adaptive cruise control, lane keeping assist), but had no experience with the Tesla Autopilot. We refrained from recruiting participants who had experience with automated driving systems, such as the Tesla's Autopilot, because participants in our previous experiment were not experienced with automated driving systems either [23] (and see [14] for a comparison of self-reported workload between automation-experienced drivers and automation-inexperienced drivers).

No incentive was provided to the participants, and all participants gave written informed consent. The study was approved by the Ethics Research Governance Office of the University of Southampton under submission ERGO number 19091.

2.2. Apparatus

The experiment was performed with a Tesla Model S 90D with Autopilot as the participants' vehicle (PV) and a Jaguar XF as a lead vehicle (LV). The LV was used for safety reasons and for creating a persistent car following task. With a forward-looking radar, forward-facing camera, and ultrasonic sensors, the Autopilot can steer, adjust speed, detect obstacles, and apply brakes automatically [32]. The Tesla Autopilot can be characterised as SAE J3016 level 2 automation (i.e., partial automation) because both steering and speed control are automated, and the driver is still expected to monitor the driving environment [33].

The Traffic-Aware Cruise Control (TACC) of the PV was set to 1, which was the closest following distance and which translates to a time headway of about 1 second. This headway corresponds to common headways in highway traffic [34, 35, 36, 37], and was sufficiently short to have a low likelihood of other cars merging in between the PV and LV.

Participants wore electrocardiography (ECG) equipment linked to LabChart 8 that captured their cardiovascular and respiratory activity. This ECG equipment consisted of the AD Instruments PowerLab 26T Teaching Series, three MLA2505 biopotential electrodes and lead wires with disposable ECG electrode patches, and the MLT1132 respiratory belt transducer. The electrodes were placed in a triangular configuration. For male participants, one electrode was placed over the xiphoid process, and two electrodes below the far ends of the collar. For female participants, one electrode was placed at the top of the sternum and two electrodes below the ribs on both sides. This gender-based distinction was mainly made for comfort purposes [see e.g., 38]. The respiratory belt was placed over the clothes around the chest.

2.3. Environment

The experiment took place on March 14–18, 2016. Participants drove on the left (slow) lanes of the British dual three-lane motorways M40, M42, and M5, for which the speed limit is 70 mph (112 km/h). Participants completed two drives during daytime outside of rush hours. The first drive was completed between entry point 14 of the M40 northbound and M5 northbound exit point 3 (Fig. 2). In the second drive, the participants drove back to the starting point. Specifically, the second drive was completed between the motorway entry point at the service stations after entry point 3 of the M5 southbound and the M40 southbound until exit point 14.

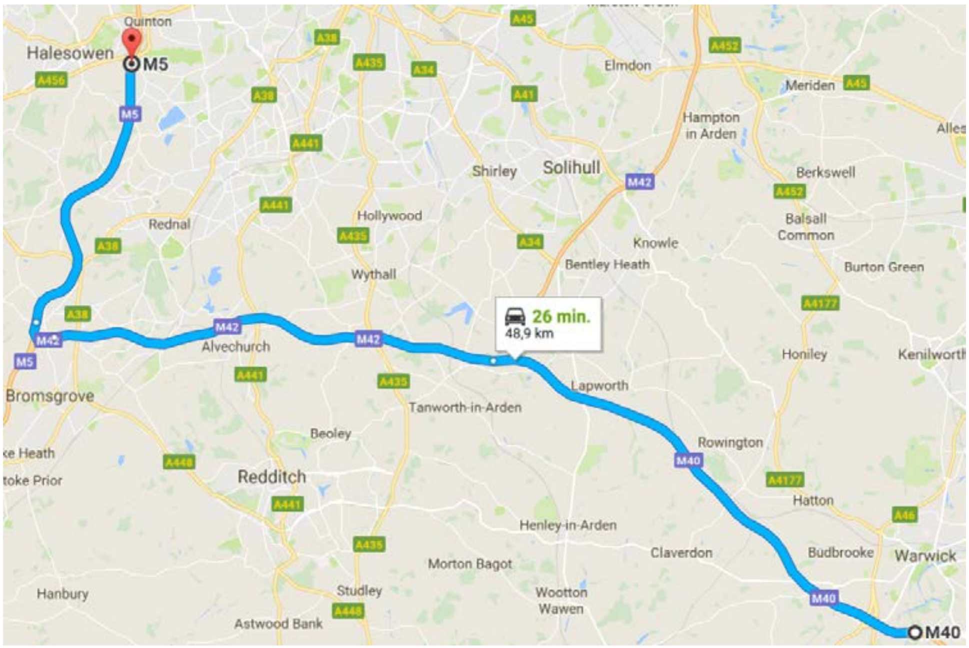


Figure 2. Map (from Google Maps) displaying the northbound route, starting at entry point 14 of the M40, and ending at exit point 3 of the M5. The southbound route went in the opposite direction, starting at entry point 3 of the M5, and ending at exit point 14 of the M40.

2.4. Procedure

All participants received a training trial and completed two drives of approximately 32 minutes each. Vigilance research has shown that detection performance exhibits a decay function with time on task [39]. Furthermore, it has been found that after 15 minutes the most substantial deterioration of detection performance has taken place (see a review by [40], reporting that “at least half of the final loss is completed within the first 15 min). Because the average driving trip in Europe and the U.S. is between 20 and 30 minutes [41, 42], it may be assumed that the present study is representative of the first exposures to a new automated driving system on public roads.

Before the experiment, the participant performed a test drive on a test track. Upon arrival at the test track site, the participant received paper instructions explaining that he/she would be driving within a highly automated platoon. Furthermore, a consent form, a demographics questionnaire, and the pre-task Dundee Stress State Questionnaire (DSSQ) were provided. After having completed these questionnaires, the participant was taken to the passenger seat of the PV and introduced to the safety driver, who was a professional driver, trained to intervene in emergencies. The safety driver sat in the passenger seat throughout the experiment for legal and safety reasons. The safety driver performed a lap on the test track and showcased the Autopilot, as well as several details of the car. After that lap, the participant and safety driver changed seats, and the participant drove the car until they were comfortable driving manually and with the Autopilot feature. Then the ECG electrodes were attached after which the participant drove to the selected motorway entry point, following the LV. After entering the motorway, the Autopilot was engaged by the participant, and the experiment started. The safety driver sat in the seat next to the participant, and verbally intervened if the participant did not act appropriately or safely (e.g., when the participant did not override the automation when he/she should). The experimenter sat in the rear seat, monitoring the equipment and making notes of events during the experiment. Before the first drive, the participant was discouraged from interacting with the safety driver or the experimenter for the duration of the experiment. Thus, the interaction between the safety driver and the experimenter was kept to a minimum.

In the occasions where another vehicle merged in between the PV and the LV, the participants were instructed by the safety driver to remain in automated mode and follow this other vehicle. However, if the gap with the LV became large, then the participants were calmly instructed to follow the LV again by overtaking the outside traffic while it was emphasised to try to remain in automated mode. An automated lane change could be performed by using the indicator stalk while holding the steering wheel. All events such as lane changes, merges, and Autopilot (dis)engagements were recorded by the experimenter using paper and pencil. Summed across the nine participants, a total of 33 and 37 lane changes (of which 16 and 21 automated) occurred for Drive 1 and 2, respectively. A manual override occurred 9 and 4 times during Drive 1 and Drive 2, respectively.

At the end of the first drive, the participants exited the motorway and stopped at a nearby parking lot. They were then provided with the post-task DSSQ and the NASA Task Load Index (TLX). Once completed, the participants followed the LV to the motorway again and performed the second drive. At the end of the second drive, the participants were again provided with the post-task DSSQ and TLX.

2.5. Independent variables

The experiment consisted of two drives, either with (DT) or without (NT) a detection task, in counterbalanced order. Specifically, five participants completed the second ‘southbound’ drive with the detection task, and four participants completed the first ‘northbound’ drive with the detection task. Without the detection task, participants had to follow the LV as their only objective. With the detection task, they also had to detect the bridges they went underneath by pressing a handheld button (Fig. 3). In the first drive, participants drove underneath 50 bridges, and during the second drive, participants drove underneath 47 bridges. Photos of the bridges are available as supplementary material.

Learning/acclimatization effects were assessed by comparing the results of the first drive with the results of the second drive.

2.6. Dependent measures

The following measures were calculated per participant for each of the two drives:

- Duration of the drive (s).
- Mean speed (km/h), recorded with a GPS application on a smartphone.
- Hit rate of the bridges (% of bridges detected). The hit rate was calculated by linking the known locations of the bridges (as retrieved from Google Maps) with the locations at the moments of button presses (recorded with a GPS application on a smartphone).
  An algorithm was written that matched bridges with the nearest button press in terms of radial distance until all bridges were assigned to a button press or no button presses were left (each button press could be assigned to one bridge only). Button presses which followed each other within ⅔ seconds (i.e., accidental double pressing of the button) were discarded. Furthermore, if the nearest button press was more than 1,250 m from the bridge, then this bridge was marked as a miss. The liberal threshold of 1,250 was used because there were several sources of inaccuracy in the locations of the button presses. Specifically, (1) The GPS signal had a limited temporal resolution (0.2 Hz, which at an average speed of 86 km/h amounts to a travelled distance of about 120 m), (2) Some participants pressed the button late (i.e., when being beneath a bridge) while others pressed the button early (i.e., when the bridge could first be seen), and (3) The GPS recording had limited accuracy (the 50th, 95th, and 99th percentile of the estimated accuracy were 24 m, 249 m, and 965 m, respectively). By definition, the miss rate equals 100% minus the hit rate [43].
- False alarm rate (% of false alarms relative to the number of bridges). A button press was considered a false alarm when after determining the hits, there were still button presses unaccounted for. Figure 4 provides an illustration of the hits and false alarms for Participant #1.
- Heart rate (bpm). The heart rate was regarded as a measure of stress [44].
- SDNN (ms), a time-domain measure of mental workload [e.g., 45, 46, 47]. The SDNN was defined as the mean of the standard deviation (SD) of all Normal to Normal peak intervals (NN) in the ECG signal per 5-min segment along the drive [*SDNN index*, see 48]. A low SDNN value is interpreted as high workload [49; see also 23]. See Figure 5 for an illustration of the calculation process.
- LF/HF ratio, a frequency-domain measure of mental workload. This spectral analysis of the NN interval calculates the power in the low-frequency (LF) 0.04−0.15 Hz range relative to the power in the high frequency (HF) 0.15−0.40 Hz range. A high LF/HF ratio is indicative of high workload [50, 51]. Both the SDNN and the LF/HF ratio were calculated from the NN intervals after a default NN artefact filter using an open-source MATLAB program [52].
- Respiratory rate (bpm). Because the respiratory belt transducer produced a noisy signal (presumably because of in-vehicle vibrations) and may contain drifts and other artefacts, the signal was filtered with a second-order Butterworth 0.1–1.0 Hz bandpass filter. This frequency range incorporates a typical human respiratory rate of 0.25 Hz. Next, the data were rank transformed to remove outliers, and subsequently, a discrete Fourier transformation was applied to retrieve the frequency with maximum amplitude (see Fig. 6 for illustration).
- DSSQ, a self-report measure of stress and fatigue [53]. In this experiment, version 1.3 of the DSSQ was used [54]. Standardized change scores for each scale of the DSSQ were calculated as follows: (post-score−pre-score)/(standard deviation of the pre-score) [55]. The scores for the three scales (Engagement, Distress, and Worry) were calculated by averaging four subscales and averaging them to result in one score for each element [based on 56, 57, 58]. Task Engagement consists of the subscales (1) Energetic Arousal, (2) Success Motivation, (3) Intrinsic Motivation, and (4) Concentration. Distress consists of (5) Tense Arousal, (6) Hedonic Tone, (7) Control & Confidence, and (8) Anger/Frustration. Finally, Worry consists of

(9) Self-Focused Attention, (10) Self-Esteem, (11) Task-Relevant Interference, and (12) Task-Irrelevant Interference. We imputed missing answers (4% of the total) using the nearest-neighbour method.
- NASA-TLX, a self-report measure to assess workload [59]. The ‘raw’ approach was used, also known as the Raw TLX (RTLX). This approach does not apply weights to the scales [60].

The mean speed, duration, heart rate, SDNN, LF/HF ratio, and respiratory rate were calculated from the moment that the participant was 200 m in front of the first bridge until 200 m after the participant passed the last bridge. SDNN was calculated as the average across six available 5-min segments.

Figure 3. Photo taken during the experiment: the participant on the right presses the handheld button when detecting a bridge. The safety driver is sitting in the left (passenger) seat. The vehicle in front is the LV the participant had to follow during the experiment.

3. Results

Table 2 presents results per individual participant. Due to the small number of participants in this study, statistical tests are not reported, as these were deemed unreliable.

Participants drove on average about 32 km per drive, at a mean speed of 86 km/h (Table 2), which is well below the speed limit of 112 km/h (the speed limit on British motorways is 70 mph, which equals 112 km/h). The difference in duration between Drive 1 and Drive 2 is caused by the fact that Drive 1 was about 4 km longer than Drive 2. This was due to the respective entry- and exit points being in different locations.

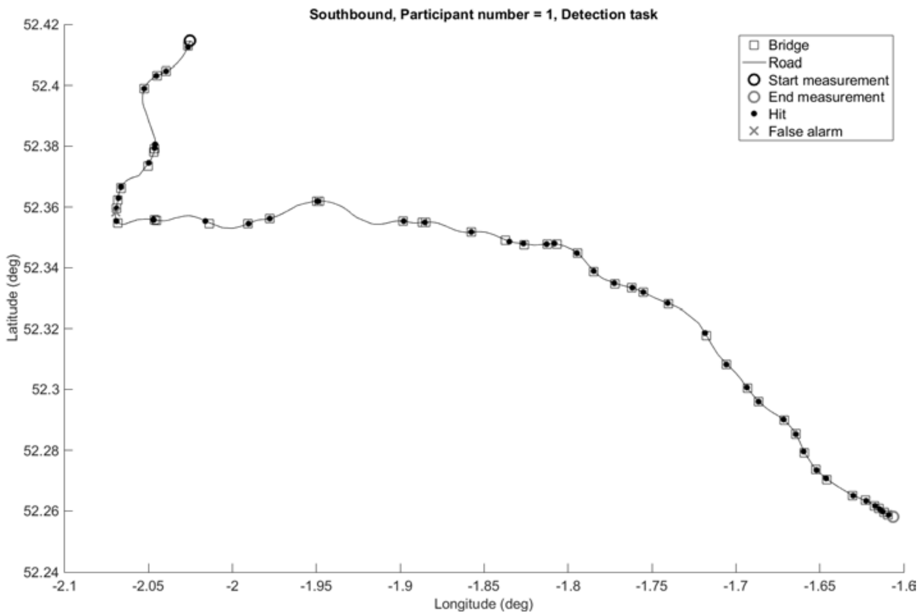
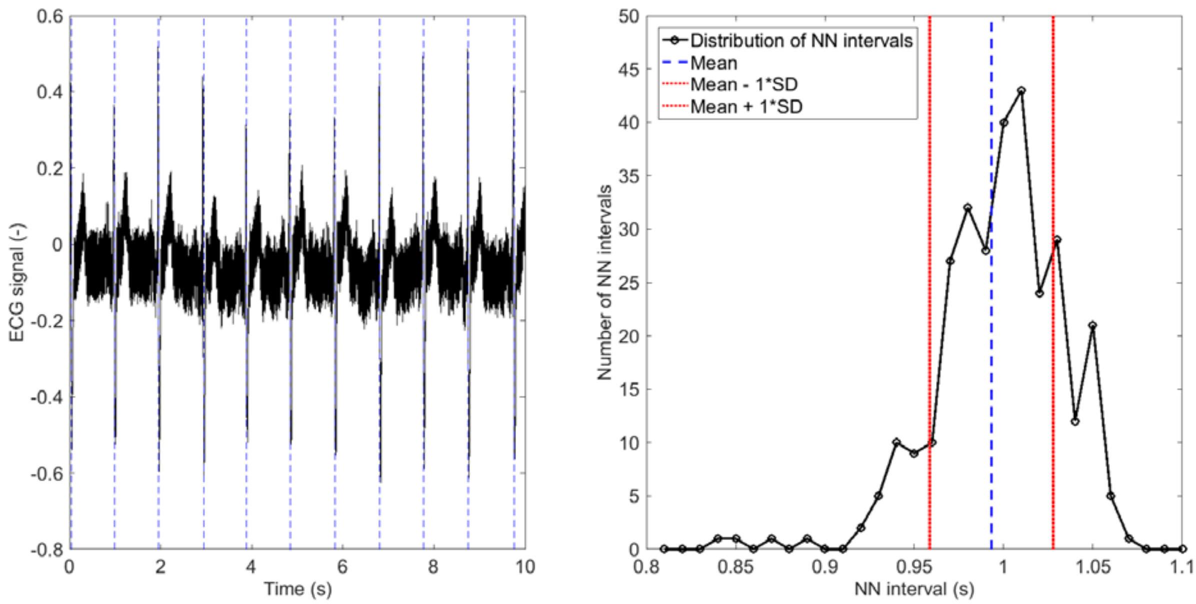
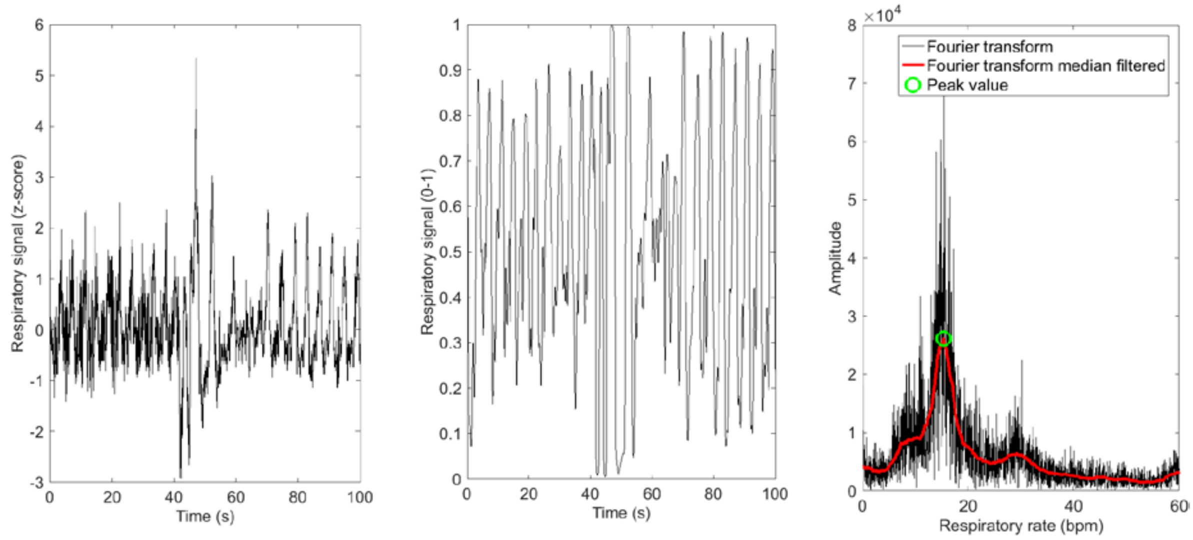


Figure 4. Illustration of GPS data and the detection task. In this case, the participant detected all 47 bridges (hit rate = 100%) and had 1 false alarm (false alarm rate = 2.1%). Where a single dot is visible for two bridges, two button presses appeared in the same GPS sample (the GPS recorded the position every 5 seconds).



A. B.

Figure 5. Illustration of the calculation of SDNN. A) ECG signal with extracted NN intervals (first 10 s of Drive 1 of Participant #1). B) Distribution of NN intervals with the mean and standard deviation of the NN intervals (SDNN = 34.6 ms; based on the first 300 s of Drive 1 of Participant #1).



A. B. C.

Figure 6. Illustration of data processing of the respiratory signal. A) *z*-transformed raw signal (first 100 s of Drive 1 of Participant #1). B) Filtered signal, rank-transformed and scaled from 0 to 1 (first 100 s of Drive 1 of Participant #1). C) Discrete Fourier transform with identified peak value (based on the entire Drive 1 of Participant #1).

The 9 participants together manually took control of the automated driving system 13 times, of which 6 were due to the Autopilot failing to anticipate on traffic merging between the LV and PV, 2 to the Autopilot following the undesired line at an exit point, 2 to the participant not trusting the Autopilot to perform correctly, 1 to an unexpected disengagement of the Autopilot, and 1 to the participant disengaging the Autopilot without apparent reason. The remaining Autopilot disengagement occurred for unknown reasons. Furthermore, lane changes were performed 29 and 41 times during the no task (NT) and detection task (DT) condition, respectively, of which 16 and 17 were manual.

Table 3 shows the results of the self-report questionnaires. Averaged across the two drives, the mean self-reported overall workload was 19% (21% in Drive 1, 16% in Drive 2; 18% for NT drives, 19% for DT drives). The mean (*SD*) per TLX item was 26% (19%) for Physical Demand, 8% (6%) for Mental Demand, 12% (9%) for Temporal Demand, 27% (29%) for Performance, 17% (13%) for Effort, and 21% (23%) for Frustration. The DSSQ results showed that participants exhibited an overall disengagement from the task, and a worriless attitude towards the task as compared to the pre-task DSSQ (i.e., the standardized change scores are smaller than 0).

Table 4 shows the results of the physiological measures. An acclimatization effect can be seen, with the heart rate being lower in Drive 2 than in Drive 1 for 8 out of 9 participants. The SDNN exhibits a negative correlation with the heart rate [see also 23]. Here, for 7 of 9 participants, SDNN was higher in Drive 2 than in Drive 1. The LF/HF ratio and respiratory rate remained relatively constant throughout the two drives.

Table 2

Descriptive results (time of day, duration, mean speed, manual overrides, manual and automated lane changes) per participant and drive number.

PP	Time of day		Duration (s)		Mean speed (km/h)		Manual overrides		Manual lane changes		Automated lane changes	
	Drive		Drive		Drive		Drive		Drive		Drive	
	1	2	1	2	1	2	1	2	1	2	1	2
1 (NT, DT)	10:52	11:55	2158	1925	80.5	82.9	1	0	2	1	0	2
2 (DT, NT)	14:29	15:18	2032	1892	85.2	84.1	0	1	1	4	5	1
3 (NT, DT)	10:43	11:46	2057	1653	85.4	96.5	2	1	1	4	1	4
4 (DT, NT)	14:40	15:20	1440	1405	85.6	86.1	1	0	2	0	2	2
5 (NT, DT)	10:33	11:28	2073	1871	83.6	85.1	2	0	0	1	3	2
6 (DT, NT)	14:41	15:33	2013	1834	86.1	87.0	0	0	2	2	1	1
7 (NT, DT)	10:25	11:18	1991	1813	87.1	87.9	0	1	1	2	2	3
8 (DT, NT)	14:34	15:34	1928	1800	88.9	88.2	2	0	2	1	1	2
9 (NT, DT)	10:40	11:42	2008	1886	86.1	84.6	1	1	3	1	1	4
Average			2033	1834	85.4	86.9	1.00	0.44	1.56	1.78	1.78	2.33

Note. NT = No Task, DT = Detection Task. Participant #4 did not complete the entire route because the batteries of the car were emptying and the car needed to be charged. This participant was excluded from the calculation of the average duration.

Table 3

Self-reported overall workload (TLX OW) and standardised change scores of self-reported stress (DSSQ) per participant and drive number

PP	TLX OW (%)		DSSQ engagement		DSSQ distress		DSSQ worry	
	Drive		Drive		Drive		Drive	
	1	2	1	2	1	2	1	2
1 (NT, DT)	12	8	-0.10	-0.21	-0.43	-0.49	-0.35	-0.57
2 (DT, NT)	18	14	-0.12	-0.18	-0.97	-1.13	0.17	-0.16
3 (NT, DT)	45	46	-1.10	-0.19	0.67	0.23	-0.78	0.14
4 (DT, NT)	17	3	-0.01	0.26	-0.82	-1.41	-2.34	-1.93
5 (NT, DT)	31	20	-0.37	-0.61	0.40	-0.16	-0.51	-0.72
6 (DT, NT)	14	14	0.18	0.46	-0.30	-0.63	-0.34	-0.32
7 (NT, DT)	9	6	-1.48	-0.52	0.46	0.26	0.49	-0.68
8 (DT, NT)	40	25	-2.60	-1.88	2.67	-0.90	0.81	1.59
9 (NT, DT)	5	7	-1.43	-0.02	1.69	1.13	-1.13	-1.69
Average	21	16	-0.78	-0.32	0.38	-0.34	-0.44	-0.48

Table 4
Cardiovascular and respiratory results per participant and drive number

PP	Heart rate (bpm)		SDNN (ms)		LF/HF ratio		Respiratory rate (bpm)	
	Drive		Drive		Drive		Drive	
	1	2	1	2	1	2	1	2
1 (NT, DT)	61.3	61.9	44.7	46.2	0.79	0.72	15.3	15.7
2 (DT, NT)	99.4	90.4	20.3	33.6	1.52	2.00	N.A.	N.A.
3 (NT, DT)	59.7	56.6	50.2	53.9	0.97	0.73	14.1	15.1
4 (DT, NT)	72.2	67.5	48.3	59.4	1.02	1.14	18.5	17.8
5 (NT, DT)	69.0	66.4	55.7	47.7	1.26	1.26	15.6	14.9
6 (DT, NT)	82.1	73.0	43.6	42.4	0.91	0.91	18.4	17.2
7 (NT, DT)	59.2	58.9	44.9	51.5	0.89	0.94	19.2	19.7
8 (DT, NT)	80.4	70.7	35.3	39.5	0.99	0.96	20.2	20.3
9 (NT, DT)	56.6	52.2	93.2	101.3	1.18	1.02	18.4	18.7
Average	71.1	66.4	48.5	52.8	1.06	1.08	17.0	16.5

Note. The respiratory rate of Participant #2 is not provided, as there was no clear peak value to be identified from the Fourier transformation.

Due to technical issues, participants #3, 4, and 6 had no button press data for 14, 16, and 10 bridges respectively. Hit rates and false alarm rates were calculated for the remaining number of bridges for these participants. The mean (*SD*) hit and false alarm rates of the bridges were 88.0% (16.0%) and 0.9% (1.5%), respectively. The lowest hit rate was 47.1%, whereas two participants had hit rates of 100% (Figure 7). The mean hit rate and mean false alarm rate correspond to a perceptual sensitivity (*d'*) of 3.52 and response bias (β) of 7.89 [61]. These results indicate that participants were well able to distinguish the bridges from the non-bridges with a conservative response strategy.

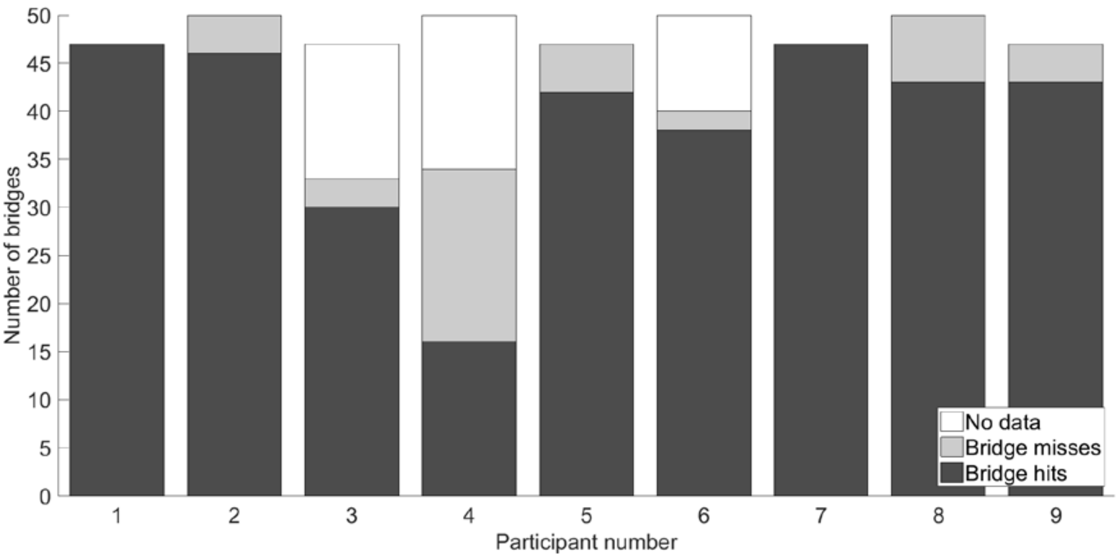


Figure 7. Number of hits and misses per participant. The number of false alarms was 1, 0, 0, 0, 1, 0, 0, 0, 2 for participants 1–9.

4. Discussion

This study aimed to measure levels of workload and stress during automated driving with the Tesla Autopilot. The literature has shown that automated driving yields low ratings of self-reported overall workload (averaging at 23%, see [9]). A previous study using the Tesla Autopilot has found overall workload scores ranging from 10% for experienced Tesla drivers on an empty highway to 43% on a city ring with drivers who had not driven in the Tesla before [Table 1; 14]. Another on-road experiment found high workload for automated driving compared to manual driving, with overall workload scores for automated driving being 42% [15]. It was unclear whether the novelty of the automation in Banks and Stanton [15] created elevated levels of workload, so they proposed extended exposure to automation, which was the purpose of the current study.

Our results of a 2 x 32 min of automated driving showed that the mean overall workload dropped from 21% in Drive 1 to 16% in Drive 2. In other words, automated driving involves a level of self-reported workload that is within the range of the workload observed in driving simulators (see Table 1, which shows a minimum overall workload of 11% and a maximum of 36%). The fact that the participants had to follow a lead vehicle may have contributed to the low overall workload by limiting their decision-making requirements. The workload was particularly low for the Physical Demand item, which may be because the participants did not have to move the pedals or steering wheel for large portions of the time. Similarly, the average heart rate and respiratory rate were close to the resting rates of a typical person [62, 63]. Thus, automated driving in the present experiment could not be considered demanding or stressful.

Compared to the detection task in Heikoop et al. [23], participants in the current study performed somewhat worse (hit rate of 95% in [23]; 88% in the present study). Furthermore, self-reported workload in the present study was not substantially different between the DT and NT conditions. It is possible that, for safety reasons or due to the presence of the safety driver sitting next to them, participants in the present study were trying harder to stay alert to the primary driving task as compared to the participants in the simulator study, thereby having a lower incentive to detect the bridges. Indeed, with the Tesla Autopilot, participants do have to remain alert due to the potential need to intervene and take manual control of the vehicle. In a survey study by Dikmen and Burns [8], 62% of Autopilot users reported that they had experienced at least one unexpected or unusual behaviour when driving in automated mode. In our experiment, the 9 participants took over manual control 9 times, for a variety of reasons, and performed 70 lane changes, of which 47% were manual. The relatively large percentage of manual lane changes may be due to the somewhat cumbersome technique required to perform an automated lane change. For an automated lane change to succeed, the driver has to press the indicator stalk while having his/her hands on the steering wheel with enough weight for the Autopilot to recognize their presence. In our experiment, this often resulted in a slight turn of the steering wheel, disengaging the Autopilot, after which the lane change had to be performed manually.

Comparing participants' self-reported engagement during a previous driving simulator study featuring similar methods [23] and another more recent simulator-based platooning study [64] with the current on-road study (Figure 8), it can be seen that participants felt relatively engaged during the on-road study. The relatively high level of engagement may be because the participants in the present study prioritized safety and tried to stay alert, as noted above. Participants of this on-road study reported relatively low levels of distress and worry compared to the participants in the simulator study. This may be because the participants in the on-road study were professional drivers who were used to driving with advanced driver assistance systems. Furthermore, the overall self-reported workload was low (a mean score of 19% on a scale from 0% to 100%). Both the on-road and simulator-based studies found that self-reported workload remained approximately constant, heart rate decreased, and SDNN increased as a function of time. Previous on-road studies into manual driving also found that the heart rate tends to decrease with time on task [65, 66]. Finally, the LF/HF ratio in the simulator and on the road were within the same range (Fig. 8).

5. Conclusions and Recommendations

This experiment complements existing research on stress and workload during automated driving (cf. Table 1), and may form a basis for more extensive research on this topic. The results point to an effect of acclimatization as demonstrated by a drop in perceived workload over time, and a decrease in heart rate. Our sample was small (N = 9); in order to acquire greater statistical power, replication studies with more participants are advised. Nonetheless, our study produced insights into the effects of workload and acclimatization to automation, and could serve as a foundation for future research into this phenomenon.

This experiment used expert drivers, and it remains to be investigated how the results translate to less trained drivers. It is likely that the participants in this study, who were experienced with various supercars and had completed advanced driver training, are more adept to novel technology and less likely to be stressed than the general population. Stapel et al. [14] found in their on-road study that people who have experience with driving in a Tesla reported a lower workload than people who had not driven a Tesla before, both during manual and automated driving.

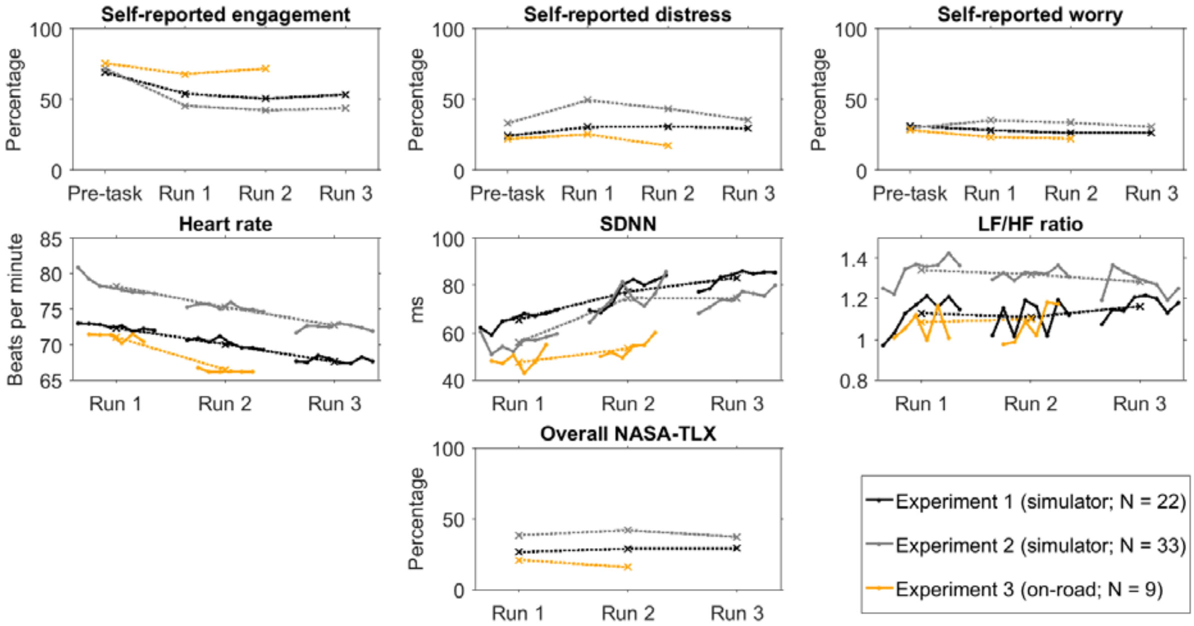


Figure 8. Comparisons of three independent experiments. Top: The Dundee Stress State Questionnaire (DSSQ). Middle: Cardiovascular measures per 5-minute interval. Bottom: Self-reported overall workload. Experiment 1: Heikoop et al. (2017) [23] (N = 22; 3 x 40 minutes of simulated driving). Experiment 2: Heikoop et al. (2018) [64] (N = 33, but N = 29 for heart rate measures]; 3 x 40 minutes of simulated driving). Experiment 3: present experiment (N = 9; 2 x 32 minutes of on-road driving). Note: The DSSQ used in Experiment 2 is the short version of the DSSQ. The DSSQ and workload scores were scaled from 0% (minimum possible) to 100% (maximum possible)

The methods used for psychophysiological measurement need consideration. Although the LF/HF ratio is regarded as a valid measure of workload [50, 51], the results of the LF/HF ratio in this experiment did not show the same time-on-task effects as the self-reported overall workload. The lack of sensitivity of the LF/HF ratio could be due to various confounding effects such as driver posture, vibrations in the vehicle, or a dependency on the heart rate itself [as also discussed by 23]. Also, the use of a smartphone-based GPS app did not allow for high accuracy data (see Section 2.6). It is recommended that future researchers use CAN bus data and differential GPS (DGPS).

The focus of this experiment was on how drivers are affected by automated driving over time. Future on-road research could include a control condition in which people drive manually [cf. 14] or include different levels of automation such as driving with ACC, or driving with ACC and steer assist [see 66]. In our experiment, control transitions were not wanted, yet occurred several times per participant. A closed-track on-road experiment [cf. 68, 69] could investigate the psychological effects of transition of control to and from automated driving in a controlled manner. Furthermore, it could be investigated whether the present findings generalize to driving tasks situations without safety driver, and with actual hazards for which manual intervention is necessary. It is possible that drivers would score better in an environment in which the target stimuli represent actual safety-critical events. Finally, it remains to be studied how our findings generalize to other driving scenarios, such as driving during rush hours.

6. Acknowledgements

This experiment has been performed in collaboration with Jaguar Land Rover. The authors are involved in the Marie Curie Initial Training Network (ITN): HFAuto (PITN-GA-2013-605817).

References

[1] Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM, 59*, 70–77. https://doi.org/10.1145/2830565

[2] Stanton, N. A., Young, M. and McCaulder, B. (1997). Drive-by-wire: the case of mental workload and the ability of the driver to reclaim control. *Safety Science, 27*(2-3), 149–159.

[3] Hancock, P. A. (2015). Automobility: The coming use of fully-automated on-road vehicles. *Proceedings of the IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 137–139.

[4] Haslbeck, A. & Hoermann, H.-J. (2016). Flying the needles: flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, *58*, 533–545.
[5] Kaber, D. K., & Endsley, M. R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress, 16*, 126–131.
[6] Metzger, U., & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 43*, 519–528.
[7] Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. *Safety Science*, *24*(1) 35–49.
[8] Dikmen, M., & Burns, C. M. (2016). Autonomous driving in the real world: Experiences with Tesla Autopilot and Summon. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16)*, 225–228, October 24–26, Ann Arbor, MI. https://doi.org/10.1145/3003715.3005465
[9] De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 196–217. https://doi.org/10.1016/j.trf.2014.06.016
[10] Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, https://doi.org/10.1177/1555343417695197
[11] Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F*, *35*, 75–84. https://doi.org/10.1016/j.trf.2015.10.005
[12] Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT press.
[13] Eriksson, A., Banks, V. A., & Stanton, N. A. (2017). Transition to manual: Comparing simulator with on-road control transitions. *Accident Analysis & Prevention, 102*, 227–234. https://doi.org/10.1016/j.aap.2017.03.011
[14] Stapel, J., Mullakka-Babu, F. A., & Happee, R. (2017). *Driver behavior and workload in an on-road automated vehicle*. Paper presented at the Road Safety & Simulation International Conference, The Hague, The Netherlands.
[15] Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics, 53*, 389–395. https://doi.org/10.1016/j.apergo.2015.06.020
[16] McDowell, K., Nunez, P., Hutchins, S., & Metcalfe, J. S. (2008). Secure mobility and the autonomous driver. *IEEE Transactions on robotics, 24*, 688–697. https://doi.org/10.1109/TRO.2008.924261
[17] Davis, J., Animashaun, A., Schoenherr, E., & McDowell, K. (2008). Evaluation of semi-autonomous convoy driving. *Journal of Field Robotics, 25*, 880–897. https://doi.org/10.1002/rob.20263
[18] Manawadu, U., Ishikawa, M., Kamezaki, M., & Sugano, S. (2015). Analysis of individual driving experience in autonomous and human-driven vehicles using a driving simulator. *Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, 299–304, July 7–11, Busan, Korea. https://doi.org/10.1109/AIM.2015.7222548
[19] Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016). Assisting drivers with ambient take-over requests in Highly Automated Driving. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16)*, 237–244, October 24–26, Ann Arbor, MI. https://doi.org/10.1145/3003715.3005409
[20] Damböck, D., Weißgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance. *Proceedings of the 16th International IEEE Annual Conference on Intelligent TransportationSystems, The Hague, The Netherlands* (pp. 1656–1661). https://doi.org/10.1109/itsc.2013.6728467
[21] De Winter, J. C. F., Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. *International Journal of Vehicle Design, 70*, 297–324. https://doi.org/10.1504/IJVD.2016.076736
[22] Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: Non-critical transitions to and from manual control. *Human Factors.* https://doi.org/10.1177/0018720816685832
[23] Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance, workload, stress, and fatigue: A driving simulator study. *Applied Ergonomics, 60*, 116–127. https://doi.org/10.1016/j.apergo.2016.10.016

[24] Large, D. R., Banks, V. A., Burnett, G. E., Baverstock, S. & Skrypchuk, L. (2017). Exploring the behaviour of distracted drivers during different levels of automation in driving. *Proceedings of the 5th International Conference on Driver Distraction and Inattention (DDI2017),* March 20–22, Paris, France.
[25] Petermeijer, S., Bazilinskyy, P., & De Winter, J. C. F. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics, 62,* 204–215. https://doi.org/10.1016/j.apergo.2017.02.023
[26] Petermeijer, S. M., Cieler, S., & De Winter, J. C. F. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis and Prevention, 99*, 218–227. https://doi.org/10.1016/j.aap.2016.12.001
[27] Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of Experimental Psychology: Applied*, *19*, 287–300. https://dx.doi.org/10.1037/a0034386
[28] Schwalk, M., Kalogerakis, N., & Maier, T. (2015). Driver support by a vibrotactile seat matrix – Recognition, adequacy and workload of tactile patterns in take-over scenarios during automated driving. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE2015), and the Affiliated Conferences,* 2466–2473, Las Vegas, NV. https://doi.org/10.1016/j.promfg.2015.07.507
[29] Young, M. S. (2000). *Attention, automaticity, and automation: New perspectives on mental underload and performance* (Doctoral dissertation). University of Southampton, UK.
[30] Young, M. S., & Stanton, N. A. (2004). Taking the load off: investigations of how adaptive cruise control affects mental workload. *Ergonomics*, *47*, 1014–1035. http://doi.org/10.1080/00140130410001686348
[31] Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*, *50*, 1324–1339. https://doi.org/10.1080/00140130701318855
[32] Full self-driving hardware on all cars. (2015). Retrieved from: https://www.tesla.com/en_GB/autopilot
[33] NHTSA (2017). Automatic vehicle control systems. Investigation PE 16-007. Retrieved from https://static.nhtsa.gov/odi/inv/2016/INCLA-PE16007-7876.pdf
[34] Brackstone, M., & McDonald, M. (2007). Driver headway: How close is too close on a motorway? *Ergonomics, 50*, 1183–1195. https://doi.org/10.1080/00140130701318665
[35] Neubert, L., Santen, L., Schadschneider, A., & Schreckenberg, M. (1999). Single-vehicle data of highway traffic – a statistical analysis. *Physical Review E, 60,* 6480–6490. https://doi.org/10.1103/PhysRevE.60.6480
[36] Song, M., & Wang, J-H. (2010). Studying the tailgating issue in Rhode Island and its treatment. *Proceedings of the 51st Annual Transportation Research Forum*, Arlington, VA, March 11-13.
[37] Treiber, M., Kesting, A., & Helbing, D. (2006). Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. *Physical Review E, 74*, 016123. https://doi.org/10.1103/PhysRevE.74.016123
[38] Shaffer, F. & Combatalade, D. C. (2013). Don't add or miss a beat: A guide to cleaner Heart Rate Variability recordings. *Biofeedback, 41*, 121–130. https://doi.org/10.5298/1081-5937-41.3.04
[39] Mackworth, J. F. (1964). Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Canadian Journal of Psychology, 18*, 209–223. https://doi.org/10.1037/h0083302
[40] Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors*, *16*, 339-353.
[41] McKenzie, B., & Rapino, M. (2011). *Commuting in the United States: 2009*. American Community Survey Reports, US Census Bureau. September, 2011.
[42] Pasaoglu, G., Fiorello, D., Martino, A., Zani, L., Zubaryeva, A., & Thiel, C. (2014). Travel patterns and the potential use of electric cars–Results from a direct survey in six European countries. *Technological Forecasting and Social Change, 87*, 51–59. https://doi.org/10.1016/j.techfore.2013.10.018
[43] Tanner Jr., W. P., Wilson, P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61*, 401–409. https://doi.org/10.1037/h0058700
[44] Healey, J. A., & Picard, R. W. (2004). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems, 6*, 156–166. https://doi.org/10.1109/TITS.2005.848368
[45] De Waard, D. (1996). *The measurement of drivers' mental workload*. Doctoral dissertation, University of Groningen.
[46] Jorna, P. G. A. M. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology, 34*, 237–257. https://doi.org/10.1016/0301-0511(92)90017-O

[47] Vidulich, M. A., & Tsang, P. S. (2012). Mental workload and situation awareness. In Salvendy, G. (Ed.), *Handbook of Human Factors and Ergonomics, Fourth Edition*, John Wiley & Sons, Inc.
[48] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation, 93*, 1043–1065. https://doi.org/ 10.1161/01.CIR.93.5.1043
[49] Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R., & Miyake, S. (2016). Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study. *Applied Ergonomics*, *52*, 95–103. https://doi.org/10.1016/j.apergo.2015.07.009
[50] Cinaz, B., Arnrich, B., La Marca, R.., & Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal and Ubiquitous Computing*, *17*, 229–239. https://doi.org/10.1007/s00779-011-0466-1
[51] Suriya-Prakash, M., John-Preetham, G., & Sharma, R. (2015). Is heart rate variability related to cognitive performance in visuospatial working memory? *PeerJ PrePrints*, *3*, e1711.
[52] Vollmer, M. (2015). A robust, simple and reliable measure of heart rate variability using relative RR intervals. IEEE Conference on Computing in Cardiology Conference (CinC), 609–612). https://doi.org/10.1109/CIC.2015.7410984
[53] Matthews, G. S., Szalma, J., Panganiban, A. R., Neubauer, C., & Warm, J. S. (2013). Profiling task stress with the Dundee Stress State Questionnaire, In: L. Cavalcanti & S. Azevedo (Eds.) *Psychology of Stress*, 49–91. Nova Science Publishers
[54] Matthews, G. S., Campbell, S. E., & Falconer, S. (2000). Personality, energetics and individual differences in motivation. *Proceedings of the Tenth European Conference on Personality*, Krakow, Poland, July 2000.
[55] Helton, W. S., Warm, J. S., Matthews, G., Corcoran, K. J., & Dember, W. N. (2002). Further tests of an abbreviated vigilance task: Effect of signal salience and jet aircraft noise on performance and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*, 1546–1550. https://doi.org/10.1177/154193120204601704
[56] Fairclough, S. H., & Venables, L. (2005). Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology, 71*, 100–110. https://doi.org/10.1016/biopsycho.2005.03.007
[57] Matthews, G. S. (2014). Task-induced fatigue, operator engagement and the control of attention. [PowerPoint slides].
[58] Matthews, G. S., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress and worry. *Emotion*, *2*, 315–340. https://doi.org/10.1037//1528-3542.2.4.315
[59] 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.
[60] Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, *50*, 904–908.
[61] Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*, 137–149. https://doi.org/10.3758/BF03207704
[62] American Heart Association. (2015, January). Target heart rates. Retrieved from http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/FitnessBasics/Target-Heart-Rates_UCM_434341_Article.jsp#.V8Q9T_krKWg
[63] Lindh, W. Q., Pooler, M. S., Tamparo, C. D., & Dahl, B. M. (2009). Vital signs and measurements, In: W. Q. Lindh, M. S. Pooler, C. D. Tamparo, & B. M. Dahl (Eds.) *Delmar's comprehensive medical assisting: administrative and clinical competencies*, Chapter 24, 573. Clifton Park, NY:Cengage Learning Inc.
[64] Lisper, H.-O., Laurell, H., & Stening, G. (1973). Effects of experience of the driver on heart-rate, respiration-rate, and subsidiary reaction time in a three hours continuous driving task. *Ergonomics, 16*, 501–506. https://doi.org/10.1080/00140137308924539
[65] Heikoop, D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2018, in press). Effects of mental demands on situation awareness during platooning: A driving simulator study. *Transportation Research Part F.*

[66] Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis and Prevention, 41*, 1087–1093. https://doi.org/10.1016/j.aap.2009.06.007

[67] Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 38*, 67–82. https://doi.org/10.1016/j.trf.2016.01.011

[68] Albert, M., Lange, A., Schmidt, A., Wimmer, M., & Bengler, K. (2015). Automated driving – Assessment of interaction concepts under real driving conditions. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE2015), and the Affiliated Conferences*, 2832–2839. https://doi.org/10.1016/j.promfg.2015.07.767.

[69] Omae, M., Hashimoto, N., Sugamoto, T., & Shimizu, H. (2005). Measurement of driver's reaction time to failure of steering controller during automated driving. *Review of Automotive Engineering*, *26*, 213–215.

Supplementary materials

Photos of the bridges obtained with Google Streetview.
https://www.dropbox.com/sh/bln7zbs0o1h9af7/AAAX2SFO2Z9vEfqu-dX4Wpiba?dl=0

GPS data of the bridges locations obtained with Google Maps, and GPS data of the button press locations of the participants. Locations in green are hits, red fields are misses, and ... are no data available. Locations in black are not applicable, i.e., those button presses fall outside of the experimental boundaries. False positives have been produced by participant #1 (at 52.3584, -2.0695), #5 (at 52.2675, -1.6380), and #9 (at 52.3056, -1.7004 and 52.2615, -1.6163).

	Bridge locs SB		**Bridge locs NB**		**1 SB**		**2 NB**		**3 SB**		**4 NB**		**5 SB**		**6 NB**		**7 SB**		**8 NB**		**9 SB**	
	Lon	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**	**lon**	**lat**
1	52,41282	-2,02633	52,25879	-1,60945	52.4125	-2.0266	52.2586	-1.6094	...	...	52.2584	-1.6092	52.4124	-2.0268	52.2582	-1.6086	52.4125	-2.0267	52.2654	-1.6132	52.4126	-2.0264
2	52,40456	-2,03963	52,25957	-1,61189	52.4046	-2.0395	52.2594	-1.6120	...	...	52.2592	-1.6114	52.4042	-2.0409	52.2593	-1.6121	52.4044	-2.0402	52.2654	-1.6132		
3	52,40318	-2,04475	52,26075	-1,61493	52.4031	-2.0452	52.2606	-1.6148	...	...	52.2604	-1.6146	52.4028	-2.0458	52.2605	-1.6150	52.4028	-2.0459	52.2654	-1.6132	52.4031	-2.0453
4	52,39906	-2,0524	52,26167	-1,61742	52.3989	-2.0524	52.2616	-1.6173	...	...	52.2605	-1.6148	52.3977	-2.0531	52.2617	-1.6174	52.3991	-2.0522	52.2654	-1.6132		
5	52,37913	-2,04629	52,26348	-1,62274	52.3806	-2.0460	52.2633	-1.6228	...	...	52.2616	-1.6171	52.3781	-2.0466	52.2633	-1.6227	52.3791	-2.0463	52.2654	-1.6132	52.3789	-2.0464
6	52,3781	-2,04674	52,26512	-1,63046	52.3792	-2.0463			...	...	52.2651	-1.6302	52.3763	-2.0480	52.2650	-1.6302	52.3766	-2.0476			52.3779	-2.0467
7	52,37326	-2,05036	52,27045	-1,64569	52.3744	-2.0499			...	...			52.3742	-2.0496	52.2698	-1.6447	52.3727	-2.0511	52.2705	-1.6462	52.3761	-2.0481
8	52,36625	-2,06655	52,27372	-1,65249	52.3665	-2.0663	52.2736	-1.6526	...	...	52.2737	-1.6528	52.3646	-2.0676	52.2738	-1.6530	52.3673	-2.0651	52.2739	-1.6532		
9	52,36236	-2,06837	52,27929	-1,65933	52.3630	-2.0681	52.2792	-1.6594	...	...			52.3621	-2.0684	52.2793	-1.6597	52.3673	-2.0651	52.2794	-1.6596	52.3596	-2.0689
10	52,35961	-2,06927	52,28527	-1,66426	52.3597	-2.0691	52.2849	-1.6641	...	...	52.2855	-1.6646			52.2850	-1.6644	52.3595	-2.0694			52.3590	-2.0691
11	52,35479	-2,06825	52,2898	-1,67121	52.3553	-2.0689	52.2901	-1.6722	...	...			52.3570	-2.0696	52.2897	-1.6714	52.3542	-2.0675	52.2899	-1.6715	52.3548	-2.0681
12	52,35574	-2,04654	52,29601	-1,68656	52.3559	-2.0473	52.2959	-1.6863	...	...	52.2958	-1.6873	52.3558	-2.0463	52.2959	-1.6863	52.3558	-2.0459	52.2936	-1.6818	52.3557	-2.0461
13	52,35562	-2,04521	52,30032	-1,69307	52.3558	-2.0470	52.3000	-1.6930	...	...			52.3555	-2.0414	52.3004	-1.6934	52.3556	-2.0441	52.3003	-1.6933	52.3556	-2.0448
14	52,35456	-2,0138	52,30815	-1,70559	52.3554	-2.0162	52.3081	-1.7057	...	...	52.3081	-1.7058	52.3558	-2.0173	52.3081	-1.7058	52.3539	-2.0109				
15	52,35444	-1,9907	52,31778	-1,71767	52.3545	-1.9905	52.3175	-1.7177	52.3533	-1.9976			52.3539	-1.9932	52.3196	-1.7198	52.3543	-1.9909	52.3178	-1.7179	52.3544	-1.9917
16	52,35625	-1,97791	52,3282	-1,74042	52.3563	-1.9779	52.3278	-1.7402	52.3563	-1.9786			52.3561	-1.9805	52.3282	-1.7406	52.3620	-1.9865	52.3281	-1.7404	52.3561	-1.9786
17	52,36189	-1,94971	52,33184	-1,75523	52.3620	-1.9495	52.3324	-1.7587	52.3619	-1.9503	52.3309	-1.7524	52.3615	-1.9537	52.3318	-1.7552	52.3620	-1.9493	52.3316	-1.7549	52.3619	-1.9502
18	52,36193	-1,94811	52,33328	-1,76217	52.3620	-1.9489	52.3329	-1.7618	52.3620	-1.9486			52.3620	-1.9477	52.3319	-1.7570	52.3620	-1.9484	52.3332	-1.7623	52.3620	-1.9480
19	52,35535	-1,8985	52,33469	-1,7721	52.3553	-1.8979	52.3349	-1.7740	52.3554	-1.8981			52.3550	-1.9057	52.3351	-1.7747	52.3553	-1.8981	52.3332	-1.7623	52.3553	-1.8967
20	52,3549	-1,88701	52,33877	-1,7847	52.3550	-1.8855	52.3411	-1.7887	52.3552	-1.8968			52.3549	-1.8874	52.3388	-1.7850	52.3550	-1.8859			52.3549	-1.8870
21	52,35502	-1,88488	52,3447	-1,79442	52.3550	-1.8852	52.3444	-1.7942	52.3551	-1.8855			52.3551	-1.8853	52.3446	-1.7944	52.3550	-1.8850	52.3449	-1.7952	52.3550	-1.8849
22	52,3517	-1,85753	52,34769	-1,80647	52.3517	-1.8576	52.3475	-1.8060	52.3516	-1.8579	52.3475	-1.8060	52.3516	-1.8571	52.3478	-1.8066	52.3518	-1.8574	52.3475	-1.8071	52.3515	-1.8574
23	52,34913	-1,83725	52,34775	-1,81238	52.3487	-1.8350			52.3493	-1.8381	52.3476	-1.8124			...	...	52.3492	-1.8385	52.3476	-1.8125	52.3490	-1.8369
24	52,34765	-1,82639	52,34765	-1,82639	52.3479	-1.8264			52.3477	-1.8270			52.3477	-1.8146	...	...	52.3476	-1.8259			52.3477	-1.8267
25	52,34775	-1,81238	52,34913	-1,83725	52.3477	-1.8129	52.3490	-1.8372	52.3477	-1.8135			52.3477	-1.8125	...	...	52.3476	-1.8126	52.3487	-1.8355	52.3479	-1.8128
26	52,34769	-1,80647	52,3517	-1,85753	52.3480	-1.8081	52.3515	-1.8593	52.3476	-1.8057			52.3479	-1.8098	...	...	52.3477	-1.8070	52.3513	-1.8568	52.3477	-1.8063
27	52,3447	-1,79442	52,35502	-1,88488	52.3448	-1.7946	52.3548	-1.8842	52.3449	-1.7949			52.3411	-1.7882	...	...	52.3439	-1.7923	52.3549	-1.8846	52.3458	-1.7987
28	52,33877	-1,7847	52,3549	-1,88701	52.3390	-1.7848	52.3547	-1.8867	52.3391	-1.7853			52.3406	-1.7875	52.3548	-1.8868	52.3384	-1.7837	52.3546	-1.8869	52.3387	-1.7845
29	52,33469	-1,7721	52,35535	-1,8985	52.3349	-1.7723	52.3550	-1.9021	52.3372	-1.7815			52.3354	-1.7757	52.3552	-1.8987	52.3350	-1.7738	52.3552	-1.8978	52.3346	-1.7719
30	52,33328	-1,76217	52,36193	-1,94811	52.3333	-1.7616	52.3618	-1.9480	52.3334	-1.7628			52.3342	-1.7683	52.3619	-1.9484	52.3344	-1.7700	52.3617	-1.9482	52.3333	-1.7619
31	52,33184	-1,75523	52,36189	-1,94971	52.3319	-1.7551	52.3617	-1.9498	52.3320	-1.7556			52.3336	-1.7623	52.3618	-1.9498	52.3310	-1.7524	52.3617	-1.9498	52.3319	-1.7552
32	52,3282	-1,74042	52,35625	-1,97791	52.3283	-1.7407	52.3561	-1.9778	52.3284	-1.7409	52.3560	-1.9782	52.3297	-1.7469	52.3560	-1.9782	52.3291	-1.7421	52.3569	-1.9730	52.3283	-1.7406
33	52,31778	-1,71767	52,35444	-1,9907	52.3185	-1.7184	52.3544	-1.9904	52.3182	-1.7180	52.3544	-1.9909	52.3216	-1.7224	52.3542	-1.9908	52.3183	-1.7179	52.3543	-1.9908	52.3183	-1.7182
34	52,30815	-1,70559	52,35456	-2,0138	52.3083	-1.7057	52.3543	-2.0138	52.3085	-1.7061	52.3544	-2.0139	52.3089	-1.7072	52.3545	-2.0140	52.3090	-1.7072	52.3547	-2.0145	52.3077	-1.7047
35	52,30032	-1,69307	52,35562	-2,04521	52.3005	-1.6933	52.3554	-2.0448	52.3004	-1.6933	...	...	52.3021	-1.6954	52.3555	-2.0451	52.2994	-1.6915	52.3565	-2.0306	52.3006	-1.6931
36	52,29601	-1,68656	52,35574	-2,04654	52.2959	-1.6864	52.3556	-2.0465	52.2961	-1.6868	...	...	52.2974	-1.6886	52.3556	-2.0465	52.2972	-1.6883	52.3556	-2.0458	52.2960	-1.6866
37	52,2898	-1,67121	52,35961	-2,06927	52.2901	-1.6716	52.3614	-2.0692	52.2908	-1.6734	...	...	52.2920	-1.6767	52.3598	-2.0697	52.2901	-1.6718	52.3599	-2.0696	52.2863	-1.6654
38	52,28527	-1,66426	52,36236	-2,06837	52.2855	-1.6644	52.3624	-2.0688	52.2885	-1.6698	...	...	52.2861	-1.6650			52.2862	-1.6652	52.3612	-2.0691	52.2860	-1.6650
39	52,27929	-1,65933	52,36625	-2,06655	52.2796	-1.6595	52.3665	-2.0669	52.2797	-1.6595	...	...			...	...	52.2790	-1.6589	52.3612	-2.0691	52.2796	-1.6595
40	52,27372	-1,65249	52,37326	-2,05036	52.2735	-1.6520	52.3735	-2.0506			...	...	52.2748	-1.6539	...	...	52.2735	-1.6522	52.3735	-2.0505	52.2737	-1.6523
41	52,27045	-1,64569	52,3781	-2,04674	52.2708	-1.6462	52.3782	-2.0470			...	...	52.2719	-1.6488	...	...	52.2702	-1.6454	52.3783	-2.0469	52.2707	-1.6461
42	52,26512	-1,63046	52,37913	-2,04629	52.2652	-1.6303	52.3791	-2.0466	52.2653	-1.6311	...	...	52.2639	-1.6249	...	...	52.2651	-1.6300	52.3792	-2.0464	52.2651	-1.6303
43	52,26348	-1,62274	52,39906	-2,0524	52.2634	-1.6225	52.3990	-2.0529	52.2637	-1.6238	...	...			...	...	52.2636	-1.6234	52.3988	-2.0531	52.2633	-1.6223
44	52,26167	-1,61742	52,40318	-2,04475	52.2617	-1.6169	52.4032	-2.0456	52.2621	-1.6181	...	...	52.2633	-1.6216	52.4035	-2.0450	52.2622	-1.6181	52.4034	-2.0447	52.2619	-1.6170
45	52,26075	-1,61493	52,40456	-2,03963	52.2606	-1.6144	52.4049	-2.0399	52.2611	-1.6154	...	...	52.2608	-1.6144	52.4047	-2.0398	52.2608	-1.6142	52.4049	-2.0393	52.2606	-1.6145
46	52,25957	-1,61189	52,41282	-2,02633	52.2598	-1.6123	52.4169	-2.0243			...	...			52.4128	-2.0265	52.2608	-1.6142	52.4107	-2.0286	52.2594	-1.6114
47	52,25879	-1,60945	52,43078	-2,01769	52.2588	-1.6091	52.4305	-2.0180	52.2590	-1.6101	...	...	52.2584	-1.6080	52.4309	-2.0180	52.2608	-1.6142	52.4310	-2.0175	52.2587	-1.6088
48			52,43487	-2,01754			52.4351	-2.0176			...	...			52.4350	-2.0177						
49			52,4432	-2,018			52.4434	-2.0180			...	...			52.4432	-2.0182			52.4430	-2.0181		
50			52,44682	-2,01691			52.4467	-2.0174			...	...										