What determines drivers' speed?
A replication of three behavioural adaptation experiments in a single driving simulator study

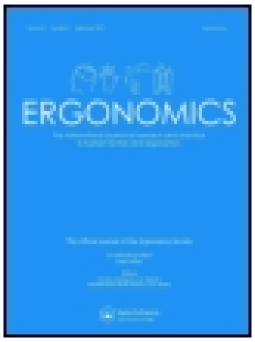
Melman, Timo; Abbink, David; van Paassen, Rene; Boer, Erwin; de Winter, Joost

DOI
10.1080/00140139.2018.1426790

Publication date
2018

Document Version
Final published version

Published in
Ergonomics: an international journal of research and practice in human factors and ergonomics



Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: http://www.tandfonline.com/loi/terg20

What determines drivers' speed? A replication of three behavioural adaptation experiments in a single driving simulator study

Timo Melman, David A. Abbink, Marinus M. van Paassen, Erwin R. Boer & Joost C. F. de Winter

To cite this article: Timo Melman, David A. Abbink, Marinus M. van Paassen, Erwin R. Boer & Joost C. F. de Winter (2018): What determines drivers' speed? A replication of three behavioural adaptation experiments in a single driving simulator study, Ergonomics, DOI: 10.1080/00140139.2018.1426790

To link to this article: https://doi.org/10.1080/00140139.2018.1426790

Accepted author version posted online: 10 Jan 2018.
Published online: 05 Feb 2018.

ERGONOMICS, 2018
https://doi.org/10.1080/00140139.2018.1426790

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

What determines drivers' speed? A replication of three behavioural adaptation experiments in a single driving simulator study

Timo Melman[a], David A. Abbink[a], Marinus M. van Paassen[b], Erwin R. Boer[a] and Joost C. F. de Winter[a]

[a]Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands; [b]Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands

ABSTRACT
We conceptually replicated three highly cited experiments on speed adaptation, by measuring drivers' experienced risk (galvanic skin response; GSR), experienced task difficulty (self-reported task effort; SRTE) and safety margins (time-to-line-crossing; TLC) in a single experiment. The three measures were compared using a nonparametric index that captures the criteria of constancy during self-paced driving and sensitivity during forced-paced driving. In a driving simulator, 24 participants completed two forced-paced and one self-paced run. Each run held four different lane width conditions. Results showed that participants drove faster on wider lanes, thus confirming the expected speed adaptation. None of the three measures offered persuasive evidence for speed adaptation because they failed either the sensitivity criterion (GSR) or the constancy criterion (TLC, SRTE). An additional measure, steering reversal rate, outperformed the other three measures regarding sensitivity and constancy, prompting a further evaluation of the role of control activity in speed adaptation.

Practitioner Summary: Results from a driving simulator experiment suggest that it is not experienced risk, experienced effort or safety margins that govern drivers' choice of speed. Rather, our findings suggest that steering reversal rate has an explanatory role in speed adaptation.

ARTICLE HISTORY
Received 19 July 2017
Accepted 3 January 2018

KEYWORDS
Behavioural adaptation; risk homeostasis; driving simulator; psychophysiology; safety margins

1. Introduction

1.1. The effects of speed on road safety

Worldwide, 1.3 million people die in traffic each year, making road traffic accidents the eighth leading cause of death (Lozano et al. 2013). Excessive speed has long been considered a primary cause of traffic accidents (Aarts and van Schagen 2006; Elvik, Christensen, and Amundsen 2004; Treat et al. 1979). An increase of speed does not only relate to an increased probability of being involved in an accident, it also aggravates the severity of accidents (Elvik, Christensen, and Amundsen 2004).

When considering the aforementioned dangers of speeding, it is disconcerting that drivers tend to drive faster when receiving technological support or when encountering a less demanding environment. For example, drivers have been found to drive with higher speeds on well-lit roads than on reference roads without lighting (Assum et al. 1999), as a result of which the attainable safety benefit (i.e. safer driving due to better visibility) is partially negated by the risks of increased driving speed. Such decreases in safety as a result of a higher adopted speed are manifestations of a more general phenomenon called behavioural adaptation (Elvik 2013; Hiraoka, Masui, and Nishikawa 2010; OECD 1990; Oviedo-Trespalacios et al. 2017; Saad 2006; Sullivan et al. 2016). Although behavioural adaptation manifested as speeding has often been found (e.g. Dragutinovic et al. 2005; Janssen and Nilsson 1993), the underlying psychological mechanisms of speed adaptation are still poorly understood (Vaa 2007).

1.2. The need for understanding behavioural adaptation

There are several reasons why the determinants of speed adaptation need to be understood. First, a good understanding is important for designing effective educational and enforcement measures. Second, knowledge about

CONTACT Timo Melman T.Melman@tudelft.nl
The supplemental data for this article is available online at https://doi.org/10.1080/00140139.2018.1426790.

speed adaptation may benefit the design of new ADAS to strike a more favourable balance between technology mediated safety improvement and motivationally inspired consumption of the offered safety. For example, we have previously shown that haptic steering feedback does not yield speed adaptation if the system disables itself when driving above a threshold speed (Melman, de Winter, and Abbink 2017). A good understanding of behavioural adaptation may allow for improvements of the algorithms and threshold settings of such technology. Third, knowledge of the determinants of speed choice may prove useful in the design of automated driving technology that behaves in a human-like (anthropomorphic) manner, rather than to adhere rigidly to a particular speed limit. It is expected that automated driving systems are better accepted if they behave anthropomorphically (e.g. Elbanhawi, Simic, and Jazar 2015; Kolekar, De Winter, and Abbink 2017; Waytz, Heafner, and Epley 2014).

1.3. Three previous experiments on speed adaptation

A large number of motivational theories of behavioural adaptation have been proposed, but the impact of three theories has been particularly large (Vaa 2007): (1) the risk homeostasis theory (Wilde 1982), (2) the task difficulty homeostasis theory (Fuller 2005) and (3) the field of safe travel theory (Gibson and Crooks 1938). These three theories, in turn, have received support from three well-cited experiments, respectively: (1) Taylor (1962), (2) Fuller, McHugh, and Pender (2008) and (3) Van Winsum and Godthelp (1996). In each of these three experiments, it is was found that an internal or external variable is sensitive to changes in driving speed, or alternatively, remains constant if drivers' change their speed. These three experiments, which are the focus of the present study, are detailed below.

1.3.1. Experiment 1: constancy of galvanic skin response (GSR) in self-paced driving (Taylor 1964)

Taylor (1962, 1964) proposed that experienced risk (i.e. anxiety level or tension) is the variable being regulated by drivers. In his research, Taylor measured the galvanic skin response (GSR, also known as electrodermal activity, as an indicator for experienced risk) of 12 participants who each drove 100 km on roads near London. Results showed that the mean GSR level per road segment did not exhibit a substantial correlation ($r = -0.04$) with the mean speed per road segment (Figure 1, left). In other words, the mean GSR was about the same regardless of whether participants were driving slowly in a busy shopping area with a high police-recorded accident rate per kilometre, or fast in a country road with a low accident rate per kilometre. These findings together with the fact that the mean GSR did correlate with driver experience (i.e. novice drivers had a higher mean GSR rate), led Taylor (1962) to conclude 'that drivers adjust their speed so that the apparent accident risk, as indicated by their rate of production of the GSR, tends to remain constant whatever the conditions'. The work of Taylor has been influential. For example, in a review, Vaa (2007) discussed Taylor's 'GSR-constancy' principle, whereas Wilde (1982, 2009) used Taylor's findings to support his risk homeostasis theory (Figure 1, left). Indeed, according to Wilde (1982), 'these findings were very instrumental in the development of the theory'.

1.3.2. Experiment 2: sensitivity of self-reported task difficulty in forced-paced driving (Fuller, McHugh, and Pender 2008)

In a more recent paper, Fuller (2005) introduced 'task difficulty homeostasis' as a key sub-goal in driving, stating that 'what drivers attempt to maintain is a level of task difficulty' (p. 461). Fuller pointed out that task difficulty is equivalent to the construct mental workload, as can be measured using self-reports such as the six-item NASA-TLX or the unidimensional Rating Scale Mental Effort (RSME). Fuller further argued that speed is the primary means for drivers to keep their experienced task difficulty at a desired level, and found support for this theory in two experiments in which participants watched videos played at different speeds (Fuller, McHugh, and Pender 2008). In one of these experiments, forty participants answered after each video 'How difficult would you find it to drive this section of road at this speed?' on a scale from 1 (extremely easy) to 7 (extremely difficult). The results of this forced-paced (i.e. non-interactive) task showed a sensitivity to different road types, and a monotonic relationship between video speed and participants' ratings of task difficulty (Figure 1, middle). More recently, Lewis-Evans and Rothengatter (2009) replicated the results of Fuller (2005) in a driving simulator, in which participants steered themselves and the results showed a similar but non-linear association between speed and reported task difficulty.

1.3.3. Experiment 3: constancy of time-to-line-crossing in self-paced driving (Van Winsum and Godthelp 1996)

In 1938, Gibson and Crooks defined a 'field of safe travel' that defines the possible paths that the car may take without being obstructed. Gibson and Crooks argued that drivers attempt to control their car to keep it in the middle of this field. In the 1930s, the field of safe travel was not operationalised, but recently, time-based safety margins have been proposed as a suitable candidate. In a review, Summala (2007) explained: 'Gibson and Crooks (1938) … demonstrate how roadway, obstacles and other road users

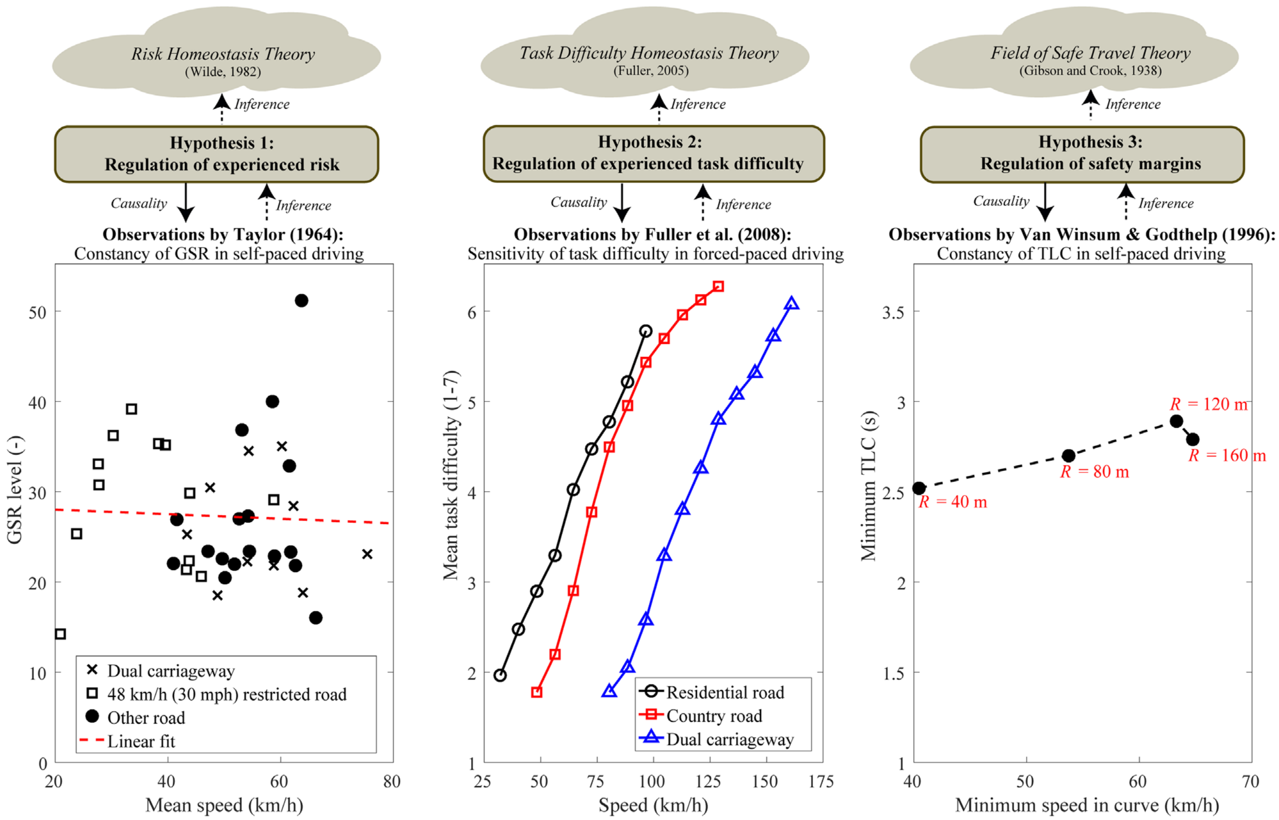


Figure 1. Relationship between speed adaptation theory (i.e. a non-operationalised set of statements), testable hypotheses and experimental observations (framework based on Meehl, 1990) Left = Experiment 1: Mean GSR level as a function of average speed on 40 road segments. Data from Taylor (1962). Middle = Experiment 2: Mean ratings of task difficulty (1 = extremely easy, 7 = extremely difficult) as a function of driving speed in videos. Data from Fuller et al. (2008, Figure 6, assuming $N = 40$). Right = Experiment 3: Mean minimum TLC as a function of mean minimum speed in curves of different radius (Van Winsum and Godthelp 1996). Speed increased substantially (60%) with increasing curve radius, while TLC showed only a moderate increase of 14% between the minimum and maximum curve radii.

modify this space – safety zone. They also implied that safety zone – and stopping distance within it – is an objectively measurable concept'. One time-based operationalisation of this field is the measure time-to-line-crossing (TLC), defined as the time it takes for the vehicle to cross the lane markers if holding the steering wheel in a steady position at the same speed (Summala 2007; van Winsum, Brookhuis, and De Waard 2000). Put differently, TLC represents the amount of time a driver has for 'error neglecting' (Godthelp 1988) or 'satisficing' (Goodrich and Boer, 2000; Summala 2007) until a corrective action is needed. Van Winsum and Godthelp (1996) showed in a driving simulator study that the minimum TLC in curves remained approximately constant with curve radius (see Figure 1, right), and they suggested 'TLC to be a regulating mechanism that determines how speed is controlled' (p. 439).

1.4. Present Study

As explained above, three influential speed adaptation theories have received corroboration from three now-classic experiments (Figure 1). In the present paper, we are not concerned with evaluating these three *theories* per se. Rather, our aim is to systematically test the three corresponding *hypotheses* in one single experiment.

The three experiments (Figure 1) were concerned with either constancy or sensitivity. We argue that a measure purporting to describe speed adaptation should meet both criteria. That is, the measure under consideration needs to remain *constant* when task demands change during *self-paced* driving (i.e. when speed adaptation is an option). This criterion was satisfied for GSR and TLC in Figure 1 left and right, respectively, as these variables remained approximately constant when the task demands (speed) changed. Second, the measure needs to be *sensitive* when task demands change during *forced-paced* driving (i.e. when speed is fixed and speed adaptation is not an option). Sensitivity was demonstrated for self-reported reported task difficulty in Figure 1 (middle), where speed adaptations were restricted.

Sensitivity alone is insufficient to validate a measure of speed adaptation, because sensitivity is uninformative about whether drivers actually use the variable to adjust their speed in self-paced conditions. Constancy alone is insufficient, as even random data or an entirely irrelevant measure would satisfy this criterion. This latter point was already recognised by Taylor (1962), who admitted that his results are 'of course consistent with the radically different assumption that the time rate of production of GSR is constant because it has nothing to do with the risk of driving'.[1]

The present study examined which of the three hypotheses [(1) regulation of experienced risk, (2) regulation of experienced task difficulty or (3) regulation of safety margins] provides the most appropriate description of speed adaptation, in terms of both sensitivity and constancy. We performed a driving simulator experiment in which participants drove on a road with cones demarcating the entire driving lane. Participants completed two forced-paced runs (i.e. fixed speed of 90 and 130 km/h, respectively) and one self-paced drive, each run at four different lane widths. We selected lane width as independent variable because lane width is a salient indicator of task demand, which, by virtue of the speed-accuracy trade-off, was expected to exhibit a strong relationship with self-paced driving speeds (De Vos, Godthelp, and Käppler 1999; Lewis-Evans and Charlton 2006; Liu, Wang, and Fu 2016; Zhai, Accot, and Woltjer 2004).

Participants reported every 20 s how much effort their current task took (cf. Fuller, McHugh, and Pender 2008), and we measured their TLC (cf. Van Winsum & Godthelp 1996) as well as their GSR (cf. Taylor 1964) while driving. Other psychophysiological measures (i.e. heart rate and heart rate variability) were recorded as well. The measures were compared with each other regarding constancy and sensitivity. Because different measures have different scale characteristics (e.g. self-reported task difficulty ranges from 0 to 10, while TLC can range from 0 to infinity on a straight road) and can be expected to respond nonlinearly to changes in lane width or speed (Lewis-Evans, De Waard, and Brookhuis 2011), we introduce a purely nonparametric method to compare the measures.

2. Methods

2.1. Participants

Twenty-four participants (17 male, 7 female) between 19 and 31 years old (M = 24.6, SD = 2.4) with normal or corrected-to-normal vision volunteered for this study. In response to the question of how often they drove in the past 12 months, one participant reported to drive every day, four drove 4–6 days a week, six drove 1–3 days per week, seven drove once a month and six drove less than once a month. Regarding mileage in the past 12 months, the most frequently selected answers were 1001–5000 km (8 respondents) and 1–1000 km (8 respondents), followed by 10,001–15,000 km (4 respondents), 5001–10,000 km (3 respondents) and 20,001–25,000 km (1 respondent). Twenty participants reported prior experience in a driving simulator, with a mean among all 24 participants of 5.3 times (SD = 10.6 times). All participant held a valid driver's licence (M = 5.8 years, SD = 2.5).

No exclusion criteria were applied regarding behaviours that are known to influence heart rate variability (HRV) and GSR, such as coffee consumption less than 2 h before the start of the experiment (11 participants), or being a smoker (2 participants) (Barutcu et al. 2005; Manzano et al. 2011; Villarejo, Zapirain, and Zorrilla 2012). However, it was not permitted to smoke or drink coffee in between the experimental sessions.

2.2. Apparatus

Participants drove in a fixed-base simulator at the Control and Simulation Department at the faculty of Aerospace Engineering, Delft University of Technology (Figure 2). Self-aligning torques of the steered front wheels were provided by a MOOG FCS ECol8000 S steering motor running at 2,500 Hz. A single-track model (heavy sedan of 1.8 m wide) was used to simulate the vehicle dynamics. The simulated vehicle had an automatic gearbox and its maximum attainable speed was 210 km/h. The environment was shown using three DLP projectors (BenQ W1080ST 1080p Full HD), together providing a horizontal and vertical field-of-view of, respectively, 180° and 40°. The images were displayed with a frame rate of 60 Hz, whereas the simulation and data logging were updated at 100 Hz. The front of the driver's car was visualised to facilitate more accurate perception of the car's position relative to the road boundaries. Constant car vibrations ('road rumble') were simulated with a seat shaker implemented in the driver's seat.

Figure 2. The fixed-based driving simulator used for the experiment.

The GSR and electrocardiographic (ECG) data were measured at 1,000 Hz using a wireless hub (Plux Wireless Biosignals S.A., Portugal). The physiological sensors were synchronised with the simulator using a 5-volt synchronisation pulse, which was initiated by the simulator at the start of each run. For the GSR measurement, one pre-gelled Ag/AgCl electrode was placed inside the hand palm and one on the side of the wrist (see also Strong 1970). The ECG local triode configuration was placed on the middle of the left chest.

2.3. Speed conditions

All participants completed three runs, each run in a different speed condition:

(1) A forced-paced condition in which the driving speed was fixed at 90 km/h (FP90).
(2) A forced-paced condition in which the driving speed was fixed at 130 km/h (FP130).
(3) A self-paced condition in which participants could adjust their speed by means of the gas and brake pedals (SP).

These conditions were counterbalanced across participants.

2.4. Lane Width Conditions and Road Environment

All participants drove each of the three runs on a single-lane 25-km long road. During each run, the participant encountered four segments of 6 km, each having a different lane width: 3.6, 2.8, 2.4 and 2.0 m. Cones were placed on the white lines to avoid that drivers would use the area outside the white lines or the hard shoulders. The lane widths allowed for a lateral deviation from the lane centre of 0.9, 0.5, 0.3 and 0.1 m, respectively, on each side of the 1.8-m-wide car before a line crossing. The lane width order was counterbalanced between runs, such that each of the 24 runs had a unique lane width order.

Each lane width segment (6 km long) consisted of five curves with 750-m inner radius and two curves with 500-m inner radius, yielding a curve/straight distance ratio of 32/68 per segment. Segments 1–4 were identical, except that the curves of Segments 1 and 3 were left/right mirrored with respect to the curves in Segments 2 and 4. A transition of lane width took place in a curve of 750-m radius. A road sign was placed 20 m before the lane-width transitions to support driver's awareness of the upcoming transition (Figure 3). Trees and cones were placed alongside the road to enhance participants' perception of speed. The cones were placed with a distance of 8 m between cones. A cone hit (defined as an incidence where the lateral error become greater than 0.9, 0.5, 0.3 or 0.1 m, depending on road width) was both visualised (i.e. red dot on the side where the car hit the cone) and made audible (a loud tone was played). No on-road obstacles and no traffic were simulated.

In order to make the driving task more challenging, the simulated car was subjected to a lateral force perturbation, applied to the car's centre of gravity. This lateral force was an unpredictable multi-sine signal consisting of five frequencies ranging from 1/15 to 1/4 Hz, and having maximum amplitude of 1000 N for the summed signal. The lateral force ensured participants needed to steer actively also on straight segments, but was not consciously noticed by most of the participants (the experimenter asked this after the experiment).

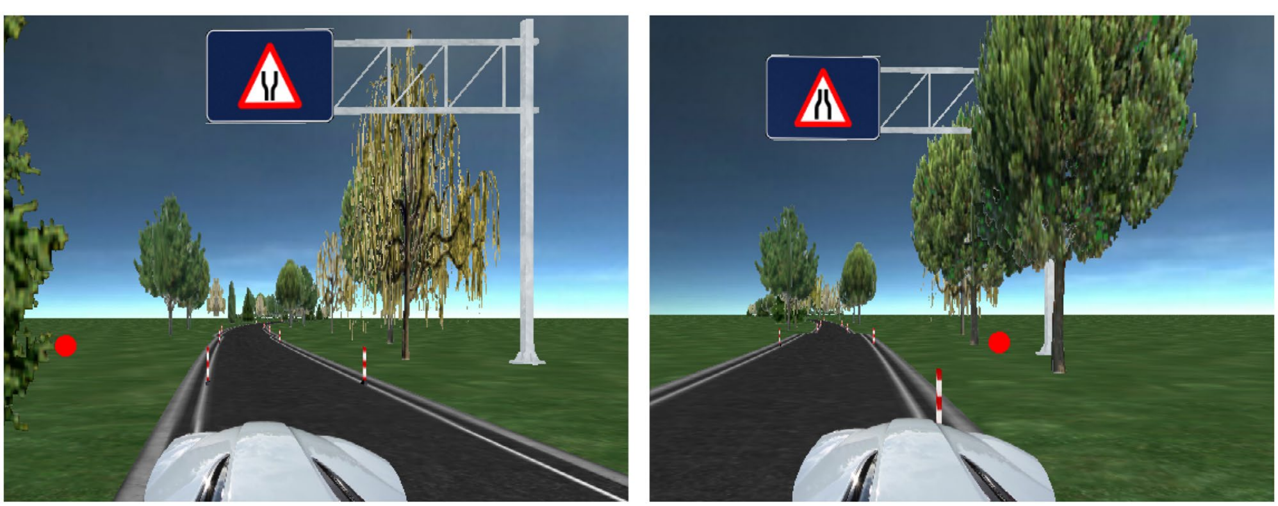


Figure 3. Simulator environment including the car front, transition signs and cone hit warning (i.e. red dot accompanied by a sound of 82 dB).

2.5. Procedures

Participants read and signed an informed consent form, which explained the purpose, instructions and procedures of the study. The consent form stated that 'the purpose of this driving-simulator study is to investigate driving behaviour, subjective experience, physiological activity, workload and comfort while driving under different task demands (i.e. lane widths)'. Participants were asked to keep both hands on the steering wheel in a ten-to-two position at all times, and were instructed to minimise the number of cone hits. The consent form further stated that every 20 s a beep would be produced to indicate that the participant had to orally report a number to the following question: 'From 0 to 10, how much effort does the current driving task takes you?', where 0 is *No effort*, 5 is *Moderate effort* and 10 is *A lot of effort*. The answers were audio-recorded and typed down by the experimenter during the experiment. The instructions (driving task and effort question) were also orally explained to ensure that all participants understood this. No speed advice was provided and participants' questions regarding speed were not answered. The speedometer was visible to the participants (see Figure 2).

Before entering the driving simulator, participants completed a questionnaire regarding their driving experience as well as a Driver Behaviour Questionnaire (DBQ) consisting of seven violation items (De Winter and Dodou 2016). A previous meta-analysis indicated that the DBQ violations scale has a moderately strong relationship ($r = 0.24$) with recorded measures of speed and speeding (De Winter, Dodou, and Stanton 2015). After completing the questionnaires, the GSR and ECG electrodes were placed and a 1-min 'rest' state was measured for the physiological variables (i.e. GSR and ECG).

Prior to the experiment, participants were familiarised with the simulator by means of a forced-paced training run followed by a self-paced training run. In the forced-paced training run, the speed was fixed at 110 km/h, the average speed of the two forced-paced test conditions. During the second training run, a beep was played every 20 s in order familiarise the participant with answering the 'effort' question. The roads of the two familiarisation runs (3.7 km each) contained the same curves and lane widths as the experimental runs. In both training runs, the four lane widths were presented in ascending order. This allowed the participants to get an indication of the range of lane widths in order to calibrate their self-reported effort ratings.

The main experiment consisted of three runs, one speed condition per run. The three speed conditions and the four lane widths were counterbalanced across participants. After each run, the participant was informed about the number of cone hits and requested to step out of the simulator for a 10-min break and to complete two questionnaires: a NASA Task Load Index (NASA-TLX) (Hart and Staveland 1988) to assess workload, the short version of the Dundee Stress State Questionnaire (DSSQ) to assess stress and fatigue (Matthews et al. 1999), and a simulator sickness item. In the latter, participants indicated whether they were feeling simulator sickness on a scale from 1 to 6 (1 = not experiencing any nausea, no sign of symptoms, 2 = arising symptoms [like a feeling in the abdomen], but no nausea, 3 = slightly nauseous, 4 = nauseous, 5 = very nauseous, retching, 6 = vomiting). The experimenters would ask the participant to leave the experiment in case that he or she provided a response of 4 or higher. The entire experiment, including placing the electrodes and completing all questionnaires, took approximately 1.5 h per participant.

2.6. Dependent measures

2.6.1. Measures per lane width

The data corresponding to the first 500 m and last 400 m of each lane width segment of 6 km were discarded in order to exclude transition effects (i.e. accelerations and decelerations) between lane widths. The following measures were calculated per lane width across 5.1 km of driving per segment.

2.6.1.1. Speed and accuracy

- *Mean Speed* (km/h).
- *Percentage Time Off-Road* (%). This is the percentage of time that the car drove outside the cone boundaries.
- *Mean and Maximum Absolute Lateral Error* (m). The absolute lateral error was defined as the distance between the middle of the car and the centre of the lane. The absolute lateral error and percentage time off-road are measures of lane-keeping accuracy.

2.6.1.2. Regulation of experienced effort

- *Mean Self-Reported Task Effort (SRTE)* (0–10). Participants reported every 20 s how much effort the current task takes from 0 (No effort) to 10 (A lot of effort). Note that we did not use Fuller, McHugh, and Pender's (2008) original wording ('How difficult would you find it to drive this section of road at this speed?') because (a) Fuller's specific wording does not apply to a self-paced task and (b) our observations from a pilot test suggested that participants tended to interpret the word 'difficult' in relation to the *objective* task demands (i.e. the lane width) rather than *subjective* experience. In order to better comply with Fuller's hypothesis, we used the word 'effort', which appears to be more in line with how difficult the participants subjectively experience the task at a particular moment (and see Kahneman 1973, for a treatise of the effort construct).

2.6.1.3. Regulation of safety margins

- *Median Time-to-Line-Crossing (*TLC) (s). The TLC was computed using a trigonometric method (van Winsum, Brookhuis, and De Waard 2000). TLC represents the time it would take for part of the vehicle to leave the lane under the assumption of constant speed and constant steering wheel angle. The TLC was assumed to be 0 s when driving outside the lane boundaries.
- *15th percentile of Time-to-Line-Crossing (TLC15th) (s)*. This measure represents the 15th percentile of the raw TLC values (Godthelp, Milgram, and Blaauw 1984). A low TLC15th or low median TLC means that drivers adopted small safety margins.

2.6.1.4. Regulation of experienced risk

- *Mean Galvanic Skin Response (GSR) (µS)*. The raw GSR signal from the left and right hands was averaged. This averaged signal was filtered using a low-pass filter (cut-off frequency of 5 Hz) to reduce extraneous noise.
- *Mean GSR Rate (µS/min)*. The rate was obtained by subtracting two adjacent sampling points of the combined mean GSR signal (explained above), taking the absolute value, and dividing this by the timestep in minutes (cf. Taylor 1964). The mean GSR may be regarded as a measure of the tonic level of the skin response, changing within tens of seconds to minutes. The mean GSR rate is a measure of the faster phasic response (Alberdi, Aztiria, and Basarab 2016; Figner and Murphy 2011; Nagai et al. 2004).
- *Mean Heart Rate (HR) (bpm)*.
- *SDNN (ms)*. This time-domain heart rate variability measure is defined as the standard deviation of the normal-to-normal (NN) intervals in the ECG signal. A low SDNN is indicative of high workload (Fallahi et al. 2016; Heikoop et al. 2017).
- *LF/HF Ratio*. This frequency-domain heart rate variability measure is defined as the ratio between the low frequencies and high frequencies of the NN intervals in the ECG signal, and offers information about sympathetic and parasympathetic activity (Berntson et al. 1997). An increase in the LF/HF ratio is indicative of increased workload (Hayashi et al. 2009; Hjortskov et al. 2004). The LF/HF ratio and SDNN were calculated after applying an NN artefact filter using software provided by Vollmer (2016).

2.6.1.5. Auxiliary measures

- *Steering Reversal Rate* (reversals/s). This is the frequency with which the steering wheel reversed direction. It was calculated by determining the local minima and maxima of the steering wheel angle; if the difference between two adjacent peaks was greater than 2 deg, it was counted as a reversal. The steering wheel angle was first filtered with a low-pass Butterworth filter with a cut-off frequency of 2 Hz.

2.6.2. Measures per speed condition

The following measures were calculated per speed condition.

- *NASA-TLX* (%). After each run, participants were asked to indicate their perceived workload for the entire run on six items: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration. Items were scored on a 21-point scale from *Very low* (0%) to *Very high* (100%), except for Performance, which ranged from *Perfect* (0%) to *Failure* (100%). The overall workload was calculated as the arithmetic mean of the six items (Byers, Bittner, and Hill 1989).
- *Dundee Stress State Questionnaire (DSSQ)*. The short multidimensional DSSQ is an operationalisation of stress and fatigue. Thirty statements were asked regarding engagement, distress and worry (Matthews et al. 1999). Items were scored from 0 (*Definitely false*) to 4 (*Definitely true*). The overall engagement, distress and worry scores ranged from 0 (minimum possible) to 32 (maximum possible).

2.7. Statistical analyses for assessing the effect of lane width and speed

For each dependent measure and for each of the three speed conditions, a matrix of 24 × 4 numbers was computed (24 participants x 4 lane width conditions). This matrix was rank transformed according to Conover and Iman (1981) to account for possible violations of the assumption of normality. The rank-transformed matrix, consisting of numbers from 1 to 96, was submitted to a repeated measures ANOVA with lane width as within-subject factor. Similarly, for each of the dependent measures, the scores for FP90 and FP130 were rank transformed according to Conover and Iman (1981). The resulting matrix, consisting of numbers from 1 to 48 (24 participants × 2 speed conditions), was submitted to a repeated measures ANOVA with the two speed conditions as within-subject factor.

2.8. Nonparametric index design to evaluate speed adaptation

We defined the amount of speed adaptation explained by a given measure using Kendall's coefficient of concordance (W), which ranges from 0 to 1 (Kendall and Smith 1939). A perfect measure of speed adaptation meets the following four criteria:

2.8.1. Constancy

(1) $W_{Sp} = 0$: no concordance during self-paced driving. For example, for SRTE, $W_{Sp} = 0$ means that participants rated the SRTE of the 2.0-, 2.4-, 2.8- and 3.6-m-wide lanes in no consistent order, and thus lane width had no consistent effect on SRTE.

2.8.2. Sensitivity

(2) $W_{FP90} = 1$: full concordance during forced-paced driving at 90 km/h. For example, for SRTE, $W_{FP90} = 1$ means that participants driving in the FP90 condition unanimously rated the 2.0-, 2.4-, 2.8- and 3.6-m-wide lanes in the same order. That is, for SRTE, all participants found the 2.0-m lane more effortful than the 2.4-m lane, the 2.4-m lane more effortful than the 2.8-m lane and the 2.8-m-wide lane more effortful than the 3.6-m lane.

(3) $W_{FP130} = 1$: full concordance during forced-paced driving at 130 km/h.

(4) $W_{\Delta Fp} = 1$: full concordance between FP130 and FP90. For example, for SRTE, $W_{\Delta Fp} = 1$ means that all participants regarded FP130 as more effortful than FP90.

The above four concordance values were used to calculate an overall speed adaptation (OSA) score (Equation 1), which applies equal weight to sensitivity (W_{FP130}, W_{FP90} & $W_{\Delta FP}$) and constancy (W_{SP}). *OSA* can range between −1 (i.e. poorest possible speed adaptation measure with constancy 1 and sensitivity 0) and 1 (i.e. perfect speed adaptation measure with constancy 0 and sensitivity 1). A score of 0 occurs if the measure were uncorrelated with the experimental conditions (e.g. if totally random data were measured) or if the measure were equally sensitive during SP and FP.

$$OSA = \frac{W_{FP130} + W_{FP90} + W_{\Delta FP}}{3} - W_{SP} \qquad (1)$$

3. Results

All participants finished the experiment; none of the participants responded a score of 3 (slightly nauseous) or higher for the simulator sickness item. Specifically, from 72 responses (24 participants x 3 runs), there were 68 responses 'Not experiencing any nausea', and 4 responses of 'Arising symptoms'.

3.1. Descriptive statistics and effects of lane width

Tables 1–3 show the means and standard deviations per lane width and per dependent measure, for the FP130, FP90 and SP conditions, respectively. These tables also contain the results of the repeated measures ANOVAs regarding lane width.

Tables 1–3 show that the wider the lane, the higher the mean absolute lateral error and maximum lateral error. Lane width also had strong effects on the TLC measures and on SRTE. For the five physiological measures (mean GSR, mean GSR rate, HR, SDNN, LF/HF ratio), the effect of lane width was substantially weaker. Only the effects of SDNN were statistically significant in all three speed conditions, with the 3.6-m-wide lane yielding higher SDNN (indicative of lower workload) than the 2.0-m-wide lane.

Figure 4 shows (1) the mean speed, (2) the cumulative number of cone hits, (3) the mean SRTE, (4) the mean TLC and (5) the mean GSR as a function of travelled distance. It can be seen that over the entire trajectory, participants adopted a higher mean speed for the wider lanes. Furthermore, for the three widest lanes (i.e. 3.6, 2.8 m and 2.4 m) participants had similar mean acceleration (on straight segments) and deceleration (before curved segments) patterns. For the 2.0-m-wide lane, however, participants adopted a relatively constant mean speed across the 5.1-km-long segment. Figure 4 and Tables 1–3 further show that substantially more cones were hit for the 2.0-m-wide lane than for the three wider lanes.

Figure 4 shows that GSR does not clearly differentiate between the different lane widths, nor between the three speed conditions. The TLC and SRTE, however, are both clearly sensitive to lane width, with wider lanes yielding higher TLC and lower SRTE. Furthermore, SRTE shows to be a measure of speed adaptation. To illustrate, for the narrowest road (blue lines), SRTE was *higher* for FP130 than for SP, whereas for the widest road (red lines), SRTE was *lower*. Put differently, it appears that participants in the SP condition, to some extent, homogenised their own task demands. A similar pattern is seen for the median TLC across participants (Figure 4). These speed adaptations are described in further detail in the following section.

Table 1. Means (*M*), standard deviations (SD) and results of the repeated measures ANOVA (*p*, *F*) per dependent measure and lane width, for the self-paced condition (SP).

	2.0 m lane width	2.4 m lane width	2.8 m lane width	3.6 m lane width	
Dependent measures	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*p* value, *F*(3,69)
Mean speed (km/h)	91.7 (21.1)	123.5 (15.7)	135.3 (16.9)	148.4 (18.6)	$p = 7.27e{-}24$ $F = 90.49$
Percentage time off-road (%)	15.64 (7.48)	1.69 (1.54)	0.25 (0.64)	0.31 (0.77)	$p = 2.17e{-}28$ $F = 130.65$
Mean absolute lateral error (m)	0.057 (0.012)	0.097 (0.017)	0.116 (0.027)	0.203 (0.046)	$p = 1.73e{-}36$ $F = 241.07$
Maximum absolute lateral error (m)	0.305 (0.166)	0.391 (0.092)	0.439 (0.138)	0.781 (0.358)	$p = 2.17e{-}28$ $F = 130.65$
Self-reported task effort (SRTE) (0–10)	6.99 (1.35)	3.81 (1.48)	3.30 (1.54)	2.32 (1.38)	$p = 9.50e{-}24$ $F = 89.61$
Median TLC (s)	1.24 (0.46)	2.15 (0.43)	2.57 (0.60)	2.96 (0.67)	$p = 7.94e{-}25$ $F = 98.04$
TLC15th (s)	0.23 (0.33)	1.14 (0.22)	1.39 (0.26)	1.65 (0.30)	$p = 1.42e{-}30$ $F = 154.81$
Mean GSR (μS)	7.38 (3.34)	7.70 (3.32)	7.62 (3.21)	7.74 (3.44)	$p = 0.664$ $F = 0.530$
Mean GSR rate (μS/min)	8.73 (8.57)	8.71 (8.18)	8.75 (6.07)	9.75 (8.82)	$p = 0.550$ $F = 0.71$
Mean HR (bpm)	79.73 (11.61)	78.49 (10.63)	78.54 (11.70)	78.94 (11.11)	$p = 0.282$ $F = 1.30$
SDNN (ms)	47.43 (16.20)	51.97 (17.31)	55.92 (20.30)	55.46 (22.75)	$p = 0.047$ $F = 2.79$
LF/HF ratio	1.09 (0.39)	1.09 (0.45)	1.07 (0.40)	1.10 (0.45)	$p = 0.935$ $F = 0.142$
Steering reversal rate (reversals/s)	0.79 (0.24)	0.63 (0.19)	0.63 (0.20)	0.61 (0.18)	$p = 1.34e{-}05$ $F = 10.10$

Table 2. Means (M), standard deviations (SD) and results of the repeated measures ANOVA (p, F) per dependent measure and lane width, for the forced-paced condition at 90 km/h (FP90).

	2.0 m lane width	2.4 m lane width	2.8 m lane width	3.6 m lane width	
Dependent measures	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*p* value, *F*(3,69)
Mean speed (km/h)	90 (0)	90 (0)	90 (0)	90 (0)	
Percentage time off-road (%)	16.62 (7.37)	0.79 (1.00)	0.05 (0.15)	0.02 (0.08)	$p = 1.01e{-}28$ $F = 134.07$
Mean absolute lateral error (m)	0.059 (0.011)	0.091 (0.020)	0.112 (0.025)	0.179 (0.047)	$p = 2.10e{-}29$ $F = 141.43$
Maximum absolute lateral error (m)	0.294 (0.114)	0.331 (0.063)	0.386 (0.084)	0.547 (0.151)	$p = 1.02e{-}28$ $F = 134.07$
Self-reported task effort (SRTE) (0–10)	6.73 (1.77)	3.31 (1.39)	1.76 (1.27)	0.86 (0.97)	$p = 6.81e{-}29$ $F = 135.91$
Median TLC (s)	1.15 (0.28)	2.77 (0.37)	3.76 (0.50)	4.88 (0.60)	$p = 5.65e{-}55$ $F = 884.63$
TLC15th (s)	0.17 (0.25)	1.53 (0.20)	2.12 (0.17)	2.75 (0.22)	$p = 5.04e{-}62$ $F = 1430.17$
Mean GSR (μS)	7.86 (4.52)	7.56 (4.29)	7.70 (4.34)	7.76 (4.14)	$p = 0.200$ $F = 1.59$
Mean GSR rate (μS/min)	7.51 (6.99)	6.81 (5.85)	6.55 (4.81)	6.57 (4.22)	$p = 0.74$ $F = 0.42$
Mean HR (bpm)	77.62 (11.31)	78.22 (11.50)	76.80 (9.68)	77.22 (11.19)	$p = 0.661$ $F = 0.53$
SDNN (ms)	49.92 (23.59)	64.05 (37.46)	60.83 (28.75)	58.35 (18.50)	$p = 1.90e{-}4$ $F = 7.57$
LF/HF ratio	1.08 (0.43)	1.22 (0.45)	1.18 (0.56)	1.27 (0.42)	$p = 0.005$ $F = 4.65$
Steering reversal rate (reversals/s)	0.78 (0.21)	0.53 (0.17)	0.45 (0.15)	0.41 (0.16)	$p = 9.12e{-}21$ $F = 69.21$

3.2. *Comparing the speed adaptation measures*

Tables 1–3 and Figure 4 described the sensitivity of the measures to lane width, for each speed condition. However, to assess speed adaptation, the effect sizes for a measure need to be evaluated for the forced-paced conditions *relative* to the self-paced condition, as shown in Table 4. Here, the index of interest is the overall speed adaptation (OSA) score, as defined in Equation (1).

Table 3. Means (M), standard deviations (SD) and results of the repeated measures ANOVA (p, F) per dependent measure and lane width, for the forced-paced condition at 130 km/h (FP130).

	2.0 m lane width	2.4 m lane width	2.8 m lane width	3.6 m lane width	
Dependent measures	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*p* value, *F*(3,69)
Mean speed (km/h)	130 (0)	130 (0)	130 (0)	130 (0)	
Percentage time off-road (%)	25.99 (11.03)	2.76 (2.41)	0.28 (0.44)	0.03 (0.12)	$p = 1.47e\text{-}37$ $F = 260.62$
Mean absolute lateral error (m)	0.075 (0.022)	0.106 (0.020)	0.123 (0.023)	0.197 (0.042)	$p = 2.88e\text{-}30$ $F = 151.21$
Maximum absolute lateral error (m)	0.411 (0.199)	0.448 (0.149)	0.467 (0.106)	0.643 (0.132)	$p = 5.81e\text{-}12$ $F = 28.02$
Self-reported task effort (SRTE) (0–10)	7.77 (1.54)	4.46 (1.53)	3.26 (1.69)	1.41 (1.16)	$p = 5.83e\text{-}32$ $F = 172.10$
Median TLC (s)	0.65 (0.28)	2.00 (0.42)	2.72 (0.43)	3.53 (0.56)	$p = 5.66e\text{-}48$ $F = 545.73$
TLC15th (s)	0.02 (0.07)	0.95 (0.19)	1.43 (0.17)	1.88 (0.15)	$p = 1.60e\text{-}55$ $F = 918.50$
Mean GSR (μS)	7.56 (3.08)	7.40 (3.00)	7.35 (3.14)	7.43 (3.44)	$p = 0.787$ $F = 0.35$
Mean GSR rate (μS/min)	7.92 (5.44)	6.71 (4.15)	7.38 (5.90)	8.33 (7.16)	$p = 0.446$ $F = 0.90$
Mean HR (bpm)	79.49 (12.57)	77.50 (11.43)	77.45 (11.50)	77.14 (11.82)	$p = 0.152$ $F = 1.82$
SDNN (ms)	52.92 (25.86)	54.61 (19.53)	51.48 (19.04)	61.92 (28.92)	$p = 0.029$ $F = 3.19$
LF/HF ratio	1.09 (0.49)	1.13 (0.47)	1.05 (0.32)	1.14 (0.33)	$p = 0.265$ $F = 1.35$
Steering reversal rate (reversals/s)	0.93 (0.28)	0.67 (0.24)	0.58 (0.20)	0.50 (0.18)	$p = 8.40e\text{-}19$ $F = 57.83$

Table 4 shows that both the SRTE and median TLC are somewhat successful in describing speed adaptation, with OSA scores for these measures being greater than 0 (0.06 and 0.17, respectively). However, the GSR and GSR rate do not perform much better than random chance, with OSA values of 0.01 and 0.03, respectively. The measures of heart rate variability (SDNN, LF/HF ratio) yield OSA values greater than 0 as well (0.07 and 0.09, respectively). It is noteworthy that the highest OSA among all measures (0.43) occurred for the steering reversal rate (Table 4).

Figure 5 shows the means across participants per lane width and per speed condition for six selected measures. In agreement with Table 4 and Figure 4, SRTE is a relatively successful measure of speed adaptation (i.e. OSA > 0) as it dropped less strongly with lane width for SP than for FP. Similarly, the increase of TLC with lane width was less steep for SP than for FP. Figure 5 further shows that the GSR measures were insensitive to lane width in all three speed conditions. Overall, steering reversal rate is the most successful measure of speed adaptation, as SRR remained relatively constant in the SP condition (i.e. low W_{SP}), while being sensitive to lane width (i.e. high W_{FP130} and W_{FP90}) (Table 4).

3.3. Supplementary analyses

As shown above, the physiological measures exhibit low sensitivity to lane width, which may suggest that these measures are statistically unreliable. However, this was clearly not the case. Figure 6, for example, illustrates that the heart rate reliably reflected individual differences (ρ = 0.90). Furthermore, a temporal effect can be distinguished: the mean heart rate decreased from Run 1 ($M = 80.6$ bpm, $SD = 11.3$) to Run 3 ($M = 76.5$, $SD = 10.8$). This run effect was further analysed by submitting a 24 × 3 (24 participants × 3 speed conditions) matrix with rank-transformed numbers to a repeated-measures ANOVA, but now with the run number as within-subject factor. The results, which can be found in the supplementary materials, show that from the 17 measures, the mean HR and DSSQ Worry are significantly different between Run 1 and Run 3.

Finally, the correlation matrices in the supplementary material reveal several noteworthy patterns. In particular, participants with a higher mean HR tend to have a lower SDNN and a higher LF/HF ratio. Additionally there are strong correlations between mean GSR and mean GSR rate, as well as between DSSQ Distress and the NASA TLX (ρ between 0.61 and 0.84). In addition, driving experience (yearly mileage) correlated with the NASA TLX ($\rho = -0.35$, -0.17, -0.48 for SP, FP90 and FP130, respectively). Low correlations were found between the physiological measures and mileage ($|\rho| < 0.15$).

4. Discussion

4.1. Main findings regarding the three speed adaptation hypotheses

We aimed to test which of three regulatory hypotheses [(1) experienced risk, (2) experienced task difficulty or (3)

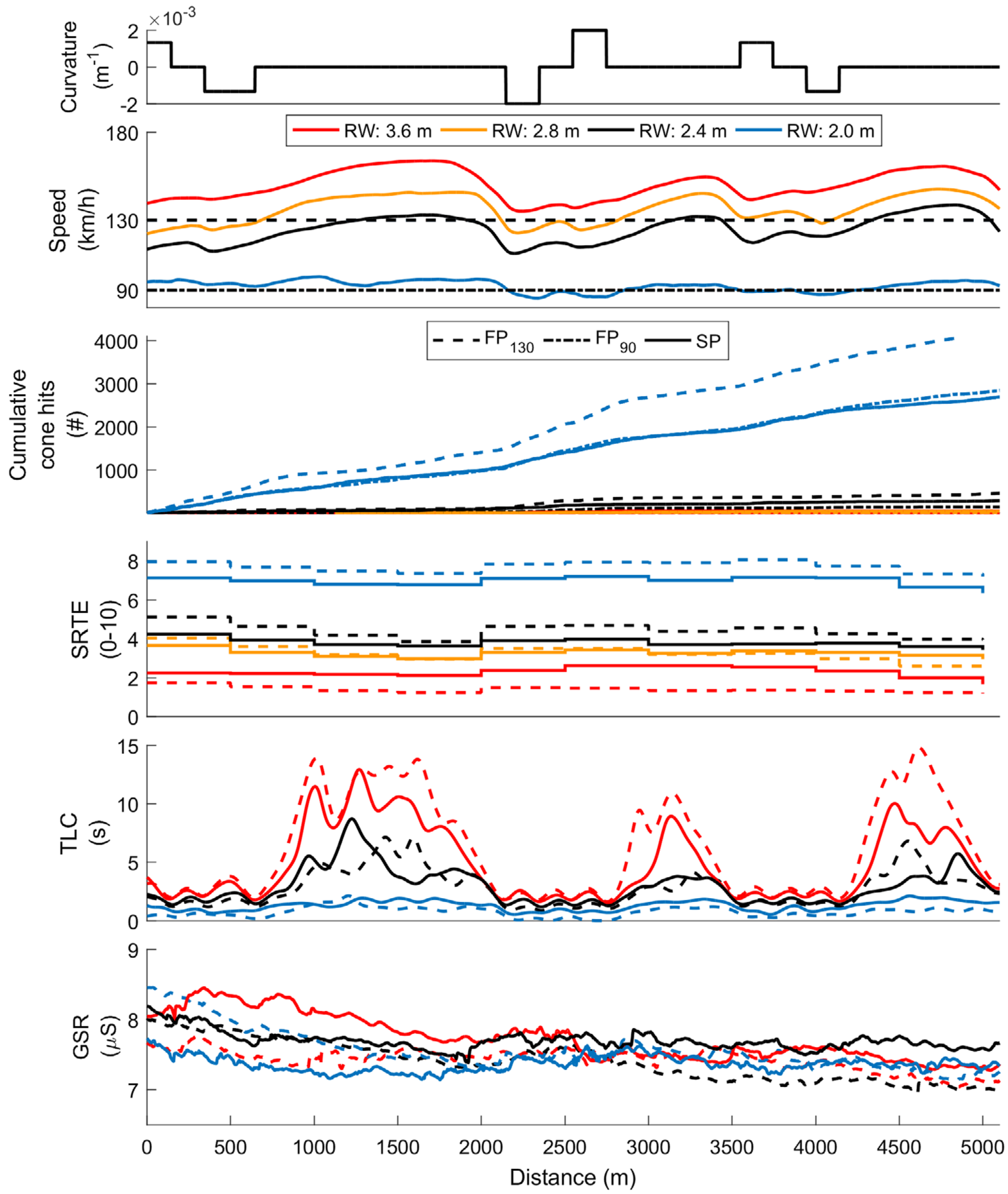


Figure 4. Selected variables as a function of travelled distance per lane width. Colours correspond to the four lane widths, and line styles correspond to the three speed conditions (for clarity, the bottom three plots do not show the FP90 condition, and the bottom two do not show the 2.8-m lane width). From top to bottom: (1) curvature (1/curve radius), (2) mean speed across participants, (3) cumulative number of cone hits summed across, (4) mean self-reported task effort (SRTE) across participants (sampled every 20 s) and (5) median TLC across participants. For visualisation purposes, the median TLC was low-pass filtered with a cut-off frequency of 0.005/m, (6) mean galvanic skin response (GSR) across participants.

safety margins] best describes the phenomenon that drivers adopt a higher speed when task demands are lowered. The three hypotheses were tested on both constancy: does the corresponding measure (i.e. GSR, SRTE, TLC) remain constant during self-paced driving?, and sensitivity: does the corresponding measure change as a function of lane width (4 lane widths) *and* imposed speed (2 fixed speeds) forced-paced driving? Previous research on this topic never tested the constancy criterion (GSR and TLC) and the sensitivity criterion (SRTE) in a single experiment, and compared the results.

Our driving simulator experiment showed that the task demand manipulation was successful in evoking speed adaptation: participants drove faster when the lane was wider. This effect, which is represented by a Kendall *W* of 0.81 (Table 3), serves as a useful confirmation that speed adaptation occurs when task demands are lowered (Lewis-Evans and Charlton 2006).

Table 4. Kendall's coefficient of concordance (W) and overall speed adaptation (OSA) per dependent measure. For the questionnaires (NASA TLX, DSSQ), the Mean (M) and Standard Deviation (SD) are reported per condition.

Dependent measure	W_{SP}	W_{FP90}	W_{FP130}	$W_{\Delta FP}$	OSA
Mean speed	0.81	–	–	1.00	–
Percentage time off-road	0.79	0.71	0.87	1.00	0.07
Mean absolute lateral error	0.94	0.88	0.88	0.56	−0.17
Maximum absolute lateral error	0.79	0.66	0.43	0.11	−0.39
Self-reported task effort (SRTE)	0.80	0.92	0.96	0.69	0.06
Median TLC	0.83	1.00	1.00	1.00	0.17
TLC15th	0.91	1.00	1.00	1.00	0.09
Mean GSR	0.00	0.00	0.00	0.03	0.01
Mean GSR rate	0.05	0.01	0.12	0.11	0.03
Mean HR	0.10	0.02	0.09	0.03	−0.05
SDNN	0.06	0.19	0.19	0.01	0.07
LF/HF ratio	0.01	0.13	0.04	0.11	0.09
Steering reversal rate	0.25	0.74	0.75	0.56	0.43
	SP	*FP90*	*FP130*		
	M (SD)	*M (SD)*	*M (SD)*	$W_{\Delta FP}$	OSA
NASA TLX (%)	52.36 (14.66)	42.15 (15.62)	58.30 (13.45)	0.84	–
DSSQ Engagement (0–32)	24.92 (4.54)	22.08 (5.01)	25.42 (4.41)	0.39	–
DSSQ Distress (0–32)	12.00 (4.61)	10.42 (4.27)	14.13 (4.82)	0.39	–
DSSQ Worry (0–32)	6.63 (5.94)	6.17 (5.57)	5.83 (5.31)	0.04	–

Note: *W* ranges between 0 and 1; OSA ranges between −1 and 1. *WSP*, *WFP90*, *WFP130*, *WΔFP* represent the Kendall's coefficient of concordance regarding (1) the effect of lane width for the self-paced condition, (2) the effect of lane width for the forced-paced 90 km/h condition, (3) the effect of lane width for forced-paced 130 km/h condition and (4) the effect of speed between the two forced-paced conditions. Furthermore, the NASA TLX and DSSQ were administered after each run, and are therefore not available per lane width.

Because the dependent measures respond non-linearly to changes in task demands (see Figure 5, for an illustration), a purely nonparametric index, called overall speed adaptation (OSA), was used. The OSA score can range between −1 and 1, where positive values mean that speed adaptation is captured by the measure; that is, the sensitivity to changes in task demand in forced paced driving conditions is greater than the sensitivity to task demand under self-paced conditions. Table 4 showed positive scores of 0.01, 0.06 and 0.17 for GSR, SRTE and TLC, respectively, which are still far from the perfect OSA = 1 score. Thus, results show that SRTE and TLC describe some speed adaptation, but none of the three tested measures provides a persuasive description of speed adaptation. The tested regulatory hypotheses failed either the criterion of sensitivity or the criterion of constancy.

4.2. Insufficient sensitivity of GSR

The mean GSR and mean GSR rate exhibited clear individual differences (as evidenced by the test–retest correlations exceeding 0.80, see Table A4 in Supplementary material), but did *not* significantly co-vary with lane width or with the imposed speed in the forced-paced conditions. This lack of sensitivity may have several causes.

First, the GSR signal exhibited large fluctuations that were uncorrelated with the experimental conditions. This suggests that GSR reflects high-frequency dynamics of the sympathetic nervous system, which may have overwhelmed the subtle changes in driver tension in response to lane width. The measurement instruments themselves may have also been a factor here. Although we did follow Taylor's (1964) method of measuring GSR on the hands, it is possible that turning of the steering wheel may have interfered with the GSR and ECG recordings (Bernardi et al. 1996; Sun et al. 2012). Thus, within-subject noise may have been an important factor reducing sensitivity. Future research could place the electrodes on other locations of the body, such as the neck (Wen et al. 2017).

Second, it is possible that GSR does not reflect changes in driver tension in simulated driving. Taylor (1962) measured drivers' GSR during real-world driving and found that GSR exhibited a strong correlation with participants' age ($\rho = -0.64$) and years since obtaining the first driver's licence ($\rho = -0.85$), but such strong correlations were not found in this study ($|\rho| < 0.15$ between participants' GSR levels and mileage).

Third, the GSR may have operated at a different time scale than the time scale with which lane width and speed were manipulated. In our study, all measures were calculated per 5.1-km segment of driving of which the first 500 m and last 400 m of each lane width were discarded to exclude transition effects. The GSR rate may have a more phasic characteristic and could therefore be especially responsive during these transition period only (e.g. Christie 1981). Future research could examine how drivers respond to transitions in task demands.

Fourth, it could be argued that GSR is not a sensitive proxy of experienced risk (e.g. Kinnear et al. 2013), and that Taylor's (1964) hypothesis, which states that drivers regulate their level of experienced risk, is false.

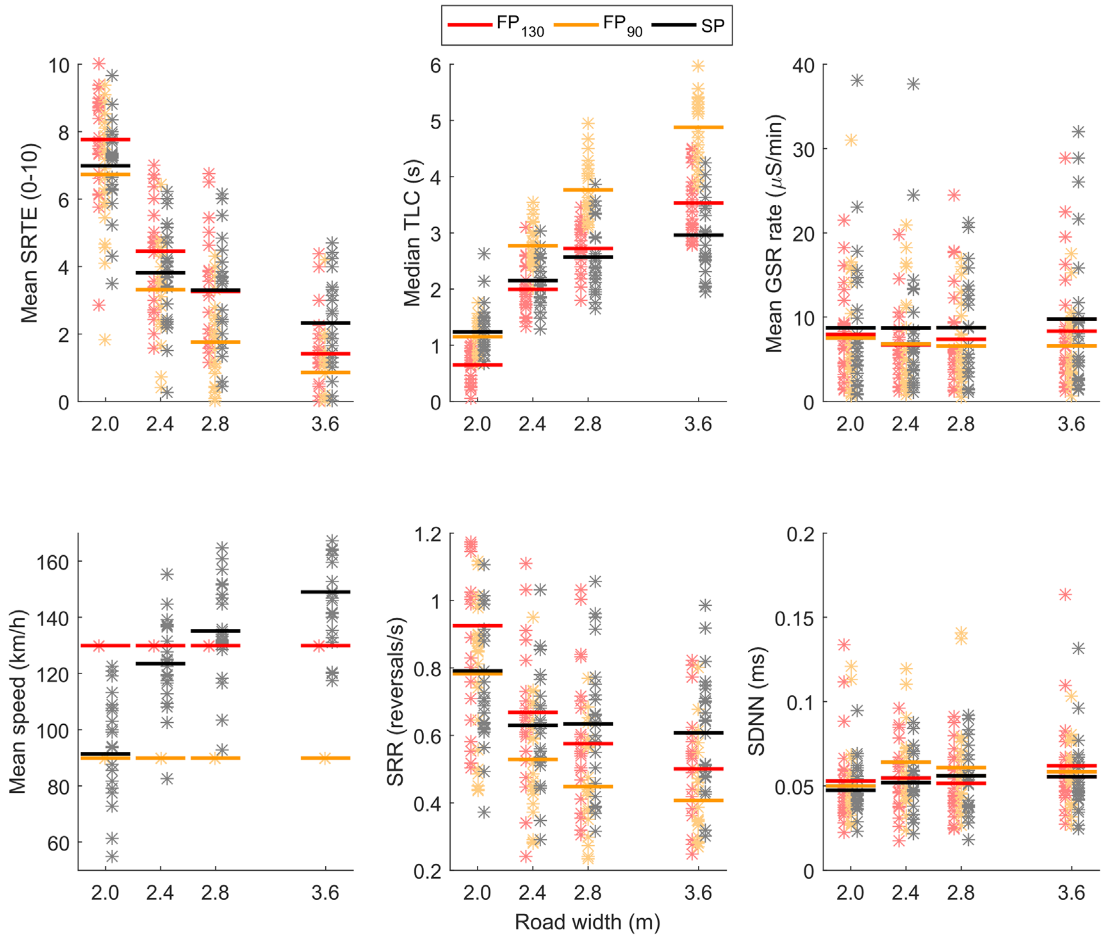


Figure 5. Scores of participants (asterisks) and means across participants (horizontal lines) per lane width (x-axis) and per speed condition (colour). Top left: self-reported task effort (SRTE); Top middle: median time-to-line-crossing (TLC); Top right: galvanic skin response rate (GSR rate); Bottom left: mean speed; Bottom middle: steering reversal rate (SRR); Bottom right: heart rate variability (SDNN).

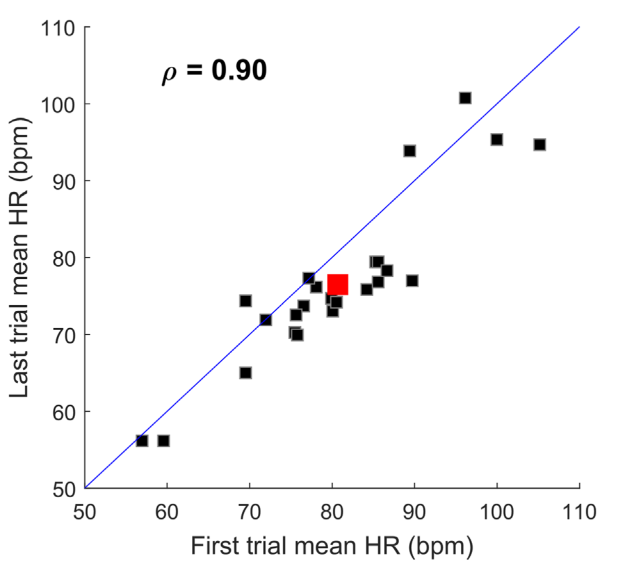


Figure 6. Mean heart rate during Run 1 versus Run 3. The markers represent values per participant (small squares) and means across 24 participants (large square).

4.3. Insufficient constancy of TLC and SRTE

Although TLC and SRTE were highly sensitive to both lane width and imposed speed, these measures were not constant during self-paced driving. Participants reported that wide lanes were less effortful to drive on (i.e. lower SRTE) than narrow lanes, even though participants drove considerably faster on the wider lanes. Here, it is possible that participants reported in congruence to what they saw (i.e. lane width itself) rather than what they subjectively experienced (i.e. experienced effort), or it is possible that Fuller, McHugh, and Pender (2008) was wrong in the sense that drivers do not regulate their experienced task difficulty.

Similarly, we found that the wider the lane, the higher the observed TLC, which may be due to the causal relationship between speed and TLC (see also the observed correlation between speed and median TLC: $\rho = -0.51$ in the Supplementary material). If maintaining the same driving path, infeasibly high speeds of 350–800 km/h (exceeding the maximum vehicle speed of 210 km/h) would have to be adopted on the widest lanes in order to acquire the same TLC as on the narrowest lane (estimated using data in Table 1). Thus, although TLC may be kept constant in some cases, such as when drivers adapt their speed to different curve radii (Van Winsum and Godthelp 1996), it failed the constancy criterion when it came to lane width.

4.4. A promising alternative measure of speed adaptation: steering reversal rate

The three behavioural adaptation hypotheses, compared in this paper, focus on subjective effort and physiological stress as well as objective risk in the form of TLC. None of the hypotheses targets objective effort. The steering reversal rate (SRR), a widely used measure of steering activity (McLean and Hoffmann 1975; Östlund et al. 2005), which may be seen as an *objective* measure of effort (Boer and Ward 2003), had the highest OSA score (0.43) of the included measures.

4.4.1. SRR yielded high sensitivity for forced-paced driving

During forced-paced driving, participants exhibited a higher SRR when the lane was narrower (i.e. W_{FP90} & W_{FP130} were high, see Table 4) and a higher SRR when the imposed speed was higher (i.e. $W_{\Delta FP}$ was high). These findings replicate early on-road research by McLean and Hoffmann (1972) which concluded that 'the proportion of high-frequency (>0.4 Hz) steering control movements increases with increasing speed and decreasing lane width, that is, increases as the driving situation becomes 'tighter' (435).

The high sensitivity of SRR to lane width (i.e. high W_{FP90} & W_{FP130}) can be explained by the fact that a larger absolute lateral error is permitted on a wider lane, and thus less frequent steering input is needed to stay in the lane. Second, a decrease of lane width is accompanied by an increase of visual saliency and thus perceptual accuracy of the vehicle state relative to the environmentally imposed constraints; that is, the distance and splay angles to the lane edges are more clearly visible when the lane is narrower (Li and Chen 2010). Indeed, steering activity is closely related to maintaining a certain vehicle state in response to perturbations such as external forces on the vehicle and perceptual inaccuracies (e.g. van Leeuwen et al. 2015). Third, the cone warnings provided salient feedback to the driver that he or she had to make a steering correction; these cone warnings occurred more frequently on the narrower lanes (see Tables 1–3).

The high sensitivity of SRR to imposed speed (i.e. high $W_{\Delta FP}$) can be explained by visual cues as well: differences in heading angles are better detectable at a higher speed due to the effects of optic flow (see Crowell and Banks 1993), thus providing incentives for steering corrections. Furthermore, a higher driving speed demands more frequent steering input due to the approximately quadratic increase in lateral displacement as a function of speed, as occurs with any vehicle (Wohl 1961).

4.4.2. SRR yielded high constancy for self-paced driving

In the self-paced condition, drivers kept a relatively constant SRR for different lane widths (i.e. W_{SP} was low). The relatively high OSA score (0.43) suggest that drivers attempt to regulate a certain control activity by means of adjusting their speed. The role of control activity in speed adaptation deserves further investigation, for example, in future experiments with a greater range of physical steering demands (e.g. sharp curves) and different task demands (e.g. higher traffic density).

4.5. Measurement considerations and temporal effects

We found that some of the dependent measures were highly correlated (see supplementary material), which indicates that a common factor may be extracted. Thus, speed adaptation may best be explained using multiple measures simultaneously. Visual scanning activity, which was not included in the present study, may be a fruitful additional measure of speed adaptation. To illustrate, it is possible that participants adapted to a decrease in task demands (i.e. increase in lane width, or a reduction in imposed speed) by engaging in extra visual scanning or by engaging in a visually distractive non-driving task. When a higher driving demand is short lived as in a slow sharp curve or a brief narrowing of a lane, drivers may temporarily increase their vigilance and posture to compensate the increased demand with increase capability. In this context, the objective measure of risk as with TLC shows an increase in risk but the perceived risk is constant because more mental effort is invested temporarily. Future research could use eye-trackers, postural sensors or brain imaging, to try to obtain a more complete picture of how drivers respond to changes in task demands.

In our study, temporal effects, in terms of the run order, were found for some measures (self-reported worry, mean heart rate). It may be argued that these temporal effects are themselves triggers of speed adaptation. On a longer time scale, it has been found that drivers' conviction rates rise in the first few years after obtaining a driver's licence (Bjørnskau and Sagberg 2005; Harrington 1972), which may be an adaptation to an increasing fearlessness while driving. Studies in which drivers' feelings and physiological measurements are recorded across multiple months are recommended to gain insight into speed adaptation during a learning process. Of course, it is also possible that the observed run order effects in our experiment simply reflect that participants became accustomed to the experimental apparatus.

4.6. Theoretical implications

We conceptually replicated three experiments that have been important in shaping extant behavioural adaptation theories (Figure 1). The fact that none of the three regulatory hypotheses convincingly described speed adaptation in our relatively simple experiment raises doubts about the validity of the three corresponding theories.

One may argue that the theories in Figure 1 are oversimplifications of actual driving and that more sophisticated theories exist nowadays. Indeed, in recent years, the theories reported in Figure 1 have been substantially revised. For example, Fuller's (2005) task difficulty homeostasis theory has been extended into a Risk Allostasis Theory by including drivers' dispositions to comply with the speed limit (Fuller 2011). Based on work of Fuller (2005), Kinnear and Helman's (2011) proposed a revised task-capability interface, a diagram with 28 blocks that are interconnected with arrows. Similar extensions also exist for Gibson and Crooks' field of safe travel (Papakostopoulos, Marmaras, and Nathanael 2017). One can argue that these sophisticated theories are more correct than the theories reported in Figure 1, because they include more factors that are known to influence driver behaviour. Although adding blocks and arrows may indeed provide a better fit to observed driving behaviour, such complexity is not necessarily theoretically convincing due to the risk of overfitting (Box 1976; Preacher 2003; Roberts and Pashler 2000). According to the well-known principle of parsimony, a theory/model should be as simple as possible, not any simpler. We recommend that researchers first determine which regulatory mechanisms occur in car driving, before devising complex models. Our findings concerning steering reversal rate calls for more research into its possible role in speed adaptation.

4.7. Experimental validity

The task demands in our driving simulator experiment were manipulated by changing the road width and imposed speed. It is possible to devise other types of task demand manipulations, including changes in traffic characteristics, weather conditions and road infrastructure (e.g. intersections, road signage). Also, our participants were mostly university students, which may hamper the generalisability of the present findings.

Another limitation is that driving in a fixed-base simulator may not sufficiently trigger driver behavioural adaptation, even though our simulator provided a large visual field of view (which improves speed perception), and incentives (task instructions, audio-visual feedback) were offered to minimise the number of cone hits. The lack of physical crash risk in a simulator could have induced a lower variety of tension levels as compared to an on-road research (e.g. see Healey and Picard 2005 for an on-road measurement of GSR). Participants in our simulator did drive considerably faster on wide lanes than on narrow lanes. On-road experiments, in which tension variability is higher, are likely to result in even greater range of speeds. Nevertheless, there are clear advantages of using a driving simulator. In particular, a simulator allows for accurate measurements of vehicle state, and for limiting the number of confounding variables. As pointed out by Taylor (1962), traffic jams or other events beyond a driver's control may prevent drivers from adopting their preferred speed (see also De Winter et al. 2007, showing that traffic turns a self-paced task into a forced-paced one). Participants in the simulator all drove in an identical environment, and could drive at a speed they preferred without being impeded.

In hindsight, we can conclude that the driving condition with the narrowest lane clearly evoked different driving behaviour than the other three lane widths, with participants barely accelerating on the straights, presumably in an attempt to minimise the number of cone hits (Figure 4). Additionally, participants experienced substantially more cone hits in the 2.0-m lane width condition than with the other three lane widths. Although our nonparametric OSA index can deal with nonlinearities, it would be worth exploring whether ceiling/floor effects or threshold effects occur at the extreme ranges of speed and road widths (see Lewis-Evans 2012 for an extensive treatise on nonlinear effects in self-reported measures during driving). Thus, whether the lane width of 2.0 m should be regarded as an outlier, or whether it is part of the full spectrum of task demand conditions, is a topic for further research. Also, participants were required to report their experienced effort every 20 s. It is possible that this secondary task itself required some effort or caused some tension and that may have manifested mostly in the 2-m-wide lane, which was already so narrow that the small amount of cognitive distraction may have been detrimental.

Lastly, this experiment was conducted with a relatively small sample size of 24 participants. Whether the SRR is truly a superior measure of speed adaptation needs to be verified in future on-road experiments with larger samples. Based on our results it is concluded that TLC and SRTE can describe some of the observed speed adaptation. The steering reversal rate shows promise in capturing speed adaptation, prompting further research into the role of conservation of control activity in car driving.

Note

1. It is noted that Taylor (1962, 1964) presented some evidence that his GSR recordings were sensitive to task demands. For example, he showed that GSR exhibited a strong negative correlation with participants' age and years of licensure (ρ = −0.64 and −0.85, respectively, based on data reported in Taylor 1962; Figures 5 and 6). Taylor (1962) also noted that participants' GSR was elevated during certain events, such as when trying to 'squeeze' their vehicle between other moving vehicles. These findings suggested that GSR is a reliable and sensitive measure of experienced risk. However, the correlation with years of licensure was based on a small sample of drivers (N = 12), while no quantitative data were provided regarding sensitivity to the external events.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is part of the research programme VIDI with project number 14127, which is financed by the Netherlands Organisation for Scientific Research (NWO). Furthermore, this work is part of the research programme VIDI with project number TTW 016.Vidi.178.047, which is financed by the Netherlands Organisation for Scientific Research (NWO).

ORCID

Joost C. F. de Winter http://orcid.org/0000-0002-1281-8200

References

Aarts, L., and I. van Schagen. 2006. "Driving Speed and the Risk of Road Crashes: A Review." *Accident Analysis and Prevention* 38 (2): 215–224. doi:10.1016/j.aap.2005.07.004.

Alberdi, A., A. Aztiria, and A. Basarab. 2016. "Towards an Automatic Early Stress Recognition System for Office Environments Based on Multimodal Measurements: A Review." *Journal of Biomedical Informatics* 59: 49–75. doi:10.1016/j.jbi.2015.11.007.

Assum, T., T. Bjørnskau, S. Fosser, and F. Sagberg. 1999. "Risk Compensation—The Case of Road Lighting." *Accident Analysis & Prevention* 31: 545–553. doi:10.1016/S0001-4575(99)00011-1.

Barutcu, I., A. M. Esen, D. Kaya, M. Turkmen, O. Karakaya, M. Melek, O.B. Esen, and Y. Basaran. 2005. "Cigarette Smoking and Heart Rate Variability: Dynamic Influence of Parasympathetic and Sympathetic Maneuvers." *Annals of Noninvasive Electrocardiology* 10: 324–329. doi:10.1111/j.1542-474X.2005.00636.x.

Bernardi, L., F. Valle, M. Coco, A. Calciati, and P. Sleight. 1996. "Physical Activity Influences Heart Rate Variability and Very-Low-Frequency Components in Holter Electrocardiograms." *Cardiovascular Research* 32 (2): 234–237.

Berntson, G. G., J. T. Bigger, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, et al. 1997. "Heart Rate Variability: Origins, Methods, and Interpretive Caveats." *Psychophysiology* 34: 623–648. doi:10.1111/j.1469-8986.1997.tb02140.x.

Bjørnskau, T., & F. Sagberg. 2005. "What Do Novice Drivers Learn during the First Months of Driving? Improved Handling Skills or Improved Road User Interaction?" In *Traffic and Transport Psychology, Theory and Application*, edited by G. Underwood, 129–140. London: Elsevier.

Boer, E. R., and N. J. Ward. 2003. "Event-Based Driver Performance Assessment." *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* 119–124. http://drivingassessment.uiowa.edu/DA2003/pdf/27_Boerformat.pdf

Box, G. E. 1976. "Science and Statistics." *Journal of the American Statistical Association* 71: 791–799.

Byers, J. C., A. C. Bittner, and S. G. Hill. 1989. "Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary?" In A. Mital, edited by *Advances in Industrial Ergonomics and Safety*, 481–485. Taylor & Francis.

Conover, W. J., and R. L. Iman. 1981. "Rank Transformations as a Bridge between Parametric and Nonparametric Statistics." *The American Statistician* 35: 124–129. doi:10.1080/00031305.1981.10479327.

Crowell, J. A., and M. S. Banks. 1993. "Perceiving Heading with Different Retinal Regions and Types of Optic Flow." *Perception & Psychophysics* 53: 325–337. doi:10.3758/BF03205187.

De Vos, A. P., J. Godthelp, and W. D. Käppler. 1999. "Subjective and Objective Assessment of Manual, Supported, and Automated Vehicle Control." In J. P. Pauwelussen, edited by *Vehicle Performance: Understanding Human Monitoring and Assessment*, 97–120. Lisse: Swets & Zeitlinger.

De Winter, J. C. F., P. A. Wieringa, J. Kuipers, J. A. Mulder, and M. Mulder. 2007. "Violations and Errors during Simulation-Based Driver Training." *Ergonomics* 50: 138–158. doi:10.1080/00140130601032721.

De Winter, J. C. F., and D. Dodou. 2016. "National Correlates of Self-Reported Traffic Violations across 41 Countries." *Personality and Individual Differences* 98: 145–152. doi:10.1016/j.paid.2016.03.091.

De Winter, J. C. F., D. Dodou, and N. A. Stanton. 2015. "A Quarter of a Century of the DBQ: Some Supplementary Notes on Its Validity with Regard to Accidents." *Ergonomics* 58: 1745–1769.

Christie, M. J. 1981. "Electrodermal Activity in the 1980s: A Review." *Journal of the Royal Society of Medicine* 74: 616–622.

Dragutinovic, N., K. A. Brookhuis, M. P. Hagenzieker, and V. A. Marchau. 2005. "Behavioural Effects of Advanced Cruise Control Use: A Meta-Analytic Approach." *European Journal of Transport and Infrastructure Research* 5: 267–280.

Elbanhawi, M., M. Simic, and R. Jazar. 2015. "In the Passenger Seat: Investigating Ride Comfort Measures in Autonomous Cars." *IEEE Intelligent Transportation Systems Magazine* 7 (3): 4–17.

Elvik, R. 2013. "20 Impact of Behavioural Adaptation on Cost-Benefit Analyses." In *Behavioural Adaptation and Road Safety: Theory, Evidence and Action*, edited by C. M. Rudin-Brown and S. L. Jamson, 371–384. Boca Raton, FL: CRC Press.

Elvik, R., P. Christensen, and A. Amundsen. 2004. *Speed and Road Accidents: An Evaluation of the Power Model (TOI Report 740)*. Oslo: Institute of Transport Economics. Retrieved from http://www.trg.dk/elvik/740-2004.pdf

Fallahi, M., M. Motamedzade, R. Heidarimoghadam, A. R. Soltanian, and S. Miyake. 2016. "Effects of Mental Workload on Physiological and Subjective Responses during Traffic

Density Monitoring: A Field Study." *Applied Ergonomics* 52: 95–103. doi:10.1016/j.apergo.2015.07.009.

Figner, B., and R. O. Murphy. 2011. "Using Skin Conductance in Judgment and Decision Making Research." In *A Handbook of Process Tracing Methods for Decision Research*, edited by M. Schulte-Mecklenbeck, A. Kuehberger and R. Ranyard, 163–184. New York, NY: Psychology Press.

Fuller, R. 2005. "Towards a General Theory of Driver Behaviour." *Accident Analysis & Prevention* 37: 461–472. doi:10.1016/j.aap.2004.11.003.

Fuller, R. 2011. "Driver Control Theory: From Task Difficulty Homeostasis to Risk Allostasis." *Handbook of Traffic Psychology*: 13–26. doi:10.1016/B978-0-12-381984-0.10002-5.

Fuller, R., C. McHugh, and S. Pender. 2008. "Task Difficulty and Risk in the Determination of Driver Behaviour." *Revue Européenne De Psychologie Appliquée/European Review of Applied Psychology* 58: 13–21. doi:10.1016/j.erap.2005.07.004.

Gibson, J. J., and L. E. Crooks. 1938. "A Theoretical Field-Analysis of Automobile-Driving." *The American Journal of Psychology* 51: 453–471. doi:10.2307/1416145.

Godthelp, H., P. Milgram, and G. J. Blaauw. 1984. "The Development of a Time-Related Measure to Describe Driving Strategy." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 26: 257–268. doi:10.1177/001872088402600302.

Godthelp, H. 1988. "The Limits of Path Error-Neglecting in Straight Lane Driving." *Ergonomics* 31: 609–619. doi:10.1080/00140138808966703.

Goodrich, M. A., and E. R. Boer. 2000. "Designing Human-centered Automation: Trade-offs in Collision Avoidance System Design." *IEEE Transactions on Intelligent Transportation Systems* 1: 40–54. doi:10.1109/6979.869020.

Harrington, D. M. 1972. "The Young Driver Follow-up Study: An Evaluation of the Role of Human Factors in the First Four Years of Driving." *Accident Analysis and Prevention* 4: 191–240. doi:10.1016/0001-4575(72)90018-8.

Hart, S. G., & L. E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In P. A. Hancock & N. Meshkati, edited by *Human Mental Workload*, vol. 52, 139–183. Amsterdam: North Holland Press.

Hayashi, T., E. Okamoto, H. Nishimura, Y. Mizuno-Matsumoto, R. Ishii, and S. Ukai. 2009. "Beta Activities in EEG Associated with Emotional Stress." *International Journal of Intelligent Computing in Medical Sciences & Image Processing* 3: 57–68. doi:10.1080/1931308X.2009.10644171.

Healey, J. A., and R. W. Picard. 2005. "Detecting Stress during Real-World Driving Tasks Using Physiological Sensors." *IEEE Transactions on Intelligent Transportation Systems* 6: 156–166. doi:10.1109/TITS.2005.848368.

Heikoop, D. D., J. C. F. de Winter, B. van Arem, and N. A. Stanton. 2017. "Effects of Platooning on Signal-Detection Performance, Workload, and Stress: A Driving Simulator Study." *Applied Ergonomics* 60: 116–127. doi:10.1016/j.apergo.2016.10.016.

Hiraoka, T., J. Masui, and S. Nishikawa. 2010. "Behavioral Adaptation to Advanced Driver-Assistance Systems." *Proceedings of SICE Annual Conference 2010*, 930–935. Taipei: IEEE.

Hjortskov, N., D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard. 2004. "The Effect of Mental Stress on Heart Rate Variability and Blood Pressure during Computer Work." *European Journal of Applied Physiology* 92: 84–89. doi:10.1007/s00421-004-1055-z.

Janssen, W. & L. Nilsson. 1993. "Behavioural Effects of Driver Support." In A. M. Parkes and S. Franzen, edited by *Driving Future Vehicles,* 147–155. London: Taylor & Francis.

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

Kendall, M. G., and B. B. Smith. 1939. "The Problem of *M* Rankings." *The Annals of Mathematical Statistics* 10: 275–287. doi:10.1214/aoms/1177732186.

Kinnear, N. A. D., and S. Helman 2011. "Updating Risk Allostasis Theory to Better Understand Behavioural Adaptation." In *Behavioural Adaptation and Road Safety*, edited by C. Rudin-Brown, and S. Jamson, 87–110. CRC Press.

Kinnear, N., S. W. Kelly, S. Stradling, and J. Thomson. 2013. "Understanding How Drivers Learn to Anticipate Risk on the Road: A Laboratory Experiment of Affective Anticipation of Road Hazards." *Accident Analysis & Prevention* 50: 1025–1033. doi:10.1016/j.aap.2012.08.008.

Kolekar, S., J. C. F. De Winter, and D. A. Abbink. 2017. "A Human-like Steering Model: Sensitive to Uncertainty in the Environment." *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, Banff, Canada.

Lewis-Evans, B. 2012. "Testing Models of Driver Behaviour." PhD thesis, University of Groningen.

Lewis-Evans, B., and S. G. Charlton. 2006. "Explicit and Implicit Processes in Behavioural Adaptation to Road Width." *Accident Analysis & Prevention* 38: 610–617. doi:10.1016/j.aap.2005.12.005.

Lewis-Evans, B., and T. Rothengatter. 2009. "Task Difficulty, Risk, Effort and Comfort in a Simulated Driving Task—Implications for Risk Allostasis Theory." *Accident Analysis & Prevention* 41: 1053–1063. doi:10.1016/j.aap.2009.06.011.

Lewis-Evans, B., D. de Waard, and K. A. Brookhuis. 2011. "Speed Maintenance under Cognitive Load–Implications for Theories of Driver Behaviour." *Accident Analysis & Prevention* 43: 1497–1507. doi:10.1016/j.aap.2011.02.030.

Li, L., and J. Chen. 2010. "Relative Contributions of Optic Flow, Bearing, and Splay Angle Information to Lane Keeping." *Journal of Vision* 10 (11): 16, 1–14, http://www.journalofvision.org/content/10/11/16. doi:10.1167/10.11.16.

Liu, S., J. Wang, and T. Fu. 2016. "Effects of Lane Width, Lane Position and Edge Shoulder Width on Driving Behavior in Underground Urban Expressways: A Driving Simulator Study." *International Journal of Environmental Research and Public Health* 13: 1010. doi:10.3390/ijerph13101010.

Lozano, R., M. Naghavi, K. Foreman, S. Lim, K. Shibuya, V. Aboyans, et al. 2013. "Global and Regional Mortality from 235 Causes of Death for 20 Age Groups in 1990 and 2010: A Systematic Analysis for the Global Burden of Disease Study 2010." *The Lancet* 380: 2095–2128. doi:10.1016/S0140-6736(12)61728-0.

Manzano, B. M., L. C. M. Vanderlei, E. M. Ramos, and D. Ramos. 2011. "Acute Effects of Smoking on Autonomic Modulation: Analysis by Poincaré Plot." *Arquivos Brasileiros De Cardiologia* 96: 154–160. doi:10.1590/S0066-782X2011005000013.

Matthews, G., L. Joyner, K. Gilliand, S. Campbell, S. Falconer, and J. Huggins. 1999. "Validation of a Comprehensive Stress State Questionnaire: Towards a State "Big Three"?" In *Personality Psychology in Europe*. Vol. 7, edited by I. Mervielde, I. J. Deary, F. de Fruyt and F. Ostendorf, 335–350. Tilburg: Tilburg University Press.

McLean, J. R., and E. R. Hoffmann. 1972. "The Effects of Lane Width on Driver Steering Control and Performance." 6th Australian Road Research Board (ARRB) Conference, Canberra 6 (3).

McLean, J. R., and E. R. Hoffmann. 1975. "Steering Reversals as a Measure of Driver Performance and Steering Task Difficulty." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 17: 248–256. Retrieved from http://www.ingentaconnect.com/content/hfes/hf/1975/00000017/00000003/art00004.

Meehl, P. E. 1990. "Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant it." *Psychological Inquiry* 1: 108–141. doi:10.1207/s15327965pli0102_1.

Melman, T., J. C. F. de Winter, and D. A. Abbink. 2017. "Does Haptic Steering Guidance Instigate Speeding? A Driving Simulator Study into Causes and Remedies." *Accident Analysis and Prevention* 98: 372–387. doi:10.1016/j.aap.2016.10.016.

Nagai, Y., H. D. Critchley, E. Featherstone, M. R. Trimble, and R. J. Dolan. 2004. "Activity in Ventromedial Prefrontal Cortex Covaries with Sympathetic Skin Conductance Level: A Physiological Account of a "Default Mode" of Brain Function." *NeuroImage* 22 (1): 243–251.

OECD. 1990. *Behaviour Adaptations to Changes in the Road Transport System*. Paris: OECD.

Östlund, J., B. Peters, B. Thorslund, J. Engström, G. Markkula, A. Keinath, et al. 2005. "Driving Performance Assessment-Methods and Metrics." http://www.aide-eu.org/pdf/sp2_deliv_new/aide_d2_2_5.pdf

Oviedo-Trespalacios, O., M. M. Haque, M. King, and S. Washington. 2017. "Effects of Road Infrastructure and Traffic Complexity in Speed Adaptation Behaviour of Distracted Drivers." *Accident Analysis & Prevention* 101: 67–77. doi:10.1016/j.aap.2017.01.018.

Papakostopoulos, V., N. Marmaras, and D. Nathanael. 2017. "The "Field of Safe Travel" Revisited: Interpreting Driving Behaviour Performance through a Holistic Approach." *Transport Reviews* 37: 695–714. doi:10.1080/01441647.2017.1289992.

Preacher, K. J. 2003. "The Role of Model Complexity in the Evaluation of Structural Equation Models." Doctoral diss., The Ohio State University.

Roberts, S., and H. Pashler. 2000. "How Persuasive is a Good Fit? A Comment on Theory Testing." *Psychological Review* 107: 358–367. doi:10.1037//0033-295x.107.2.358.

Saad, F. 2006. "Some Critical Issues When Studying Behavioural Adaptations to New Driver Support Systems." *Cognition, Technology and Work* 8: 175–181. doi:10.1007/s10111-006-0035-y.

Strong, P. 1970. *Biophysical Measurements Tektronix Measurement Series*. Beaverton, OR: Tektronix Inc.

Summala, H. 2007. "Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort through Satisficing." In *Modelling Driver Behaviour in Automotive Environments*, edited by P. C. Cacciabue, 189–207. London: Springer.

Sun, F. T., C. Kuo, H. T. Cheng, S. Buthpitiya, P. Collins, and M. Griss. 2012. "Activity-Aware Mental Stress Detection Using Physiological Sensors." *Mobile Computing, Applications, and Services* 211–230.

Sullivan, J. M., M. J. Flannagan, A. K. Pradhan, and S. Bao. 2016. *Literature Review of Behavioral Adaptations to Advanced Driver Assistance Systems*. Washington, DC: AAA Foundation for Traffic Safety.

Taylor, D. H. 1962. *A Study of Drivers' Galvanic Skin Response and the Risk of Accident* (Report No. LN/26/DHT). Road Research Laboratory: Department of Scientific and Industrial Research.

Taylor, D. H. 1964. "Drivers' Galvanic Skin Response and the Risk of Accident." *Ergonomics* 7: 439–451. doi:10.1080/00140136408930761.

Treat, J. R., N. S. Tumbas, S. T. McDonald, D. Shinar, R. D. Hume, R. E. Mayer, et al. 1979. *Tri-Level Study of the Causes of Traffic Accidents. Executive Summary* (Report No. DOT HS 805 099). Washington, DC: National Highway Traffic Safety Administration.

Vaa, T. 2007. "Modelling Driver Behaviour on Basis of Emotions and Feelings: Intelligent Transport Systems and Behavioural Adaptations." In *Modelling Driver Behaviour in Automotive Environments*, edited by P. C. Cacciabue, 208–232. London: Springer.https://doi.org/10.1007/978-1-84628-618-6_12

van Leeuwen, P. M., C. Gómez i Subils, A. Ramon Jimenez, R. Happee, and J. C. F. de Winter. 2015. "Effects of Visual Fidelity on Curve Negotiation, Gaze Behaviour and Simulator Discomfort." *Ergonomics* 58: 1347–1364. doi:10.1080/00140139.2015.1005172.

van Winsum, W., K. Brookhuis, and D. de Waard. 2000. "A Comparison of Different Ways to Approximate Time-to-Line Crossing (TLC) during Car Driving." *Accident Analysis & Prevention* 32: 47–56. doi:10.1016/S0001-4575(99)00048-2.

Van Winsum, W., and H. Godthelp. 1996. "Speed Choice and Steering Behavior in Curve Driving." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 38: 434–441. doi:10.1518/001872096778701926.

Villarejo, M. V., B. G. Zapirain, and A. M. Zorrilla. 2012. "A Stress Sensor Based on Galvanic Skin Response (GSR) Controlled by ZigBee." *Sensors* 12: 6075–6101. doi:10.3390/s120506075.

Vollmer, M. 2016. "A Robust, Simple and Reliable Measure of Heart Rate Variability Using Relative RR Intervals." *Computing in Cardiology* 42: 609–612. doi:10.1109/CIC.2015.7410984.

Waytz, A., J. Heafner, and N. Epley. 2014. "The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle." *Journal of Experimental Social Psychology* 52: 113–117. doi:10.1016/j.jesp.2014.01.005.

Wen, W., D. Tomoi, H. Yamakawa, S. Hamasaki, K. Takakusaki, Q. An, Y. Tamura, A. Yamashita, and H. Asama. 2017. "Continuous Estimation of Stress Using Physiological Signals during a Car Race." *Psychology* 08: 978–986. doi:10.4236/psych.2017.87064.

Wilde, G. J. 1982. "The Theory of Risk Homeostasis: Implications for Safety and Health." *Risk Analysis* 2: 209–225. doi:10.1111/j.1539-6924.1982.tb01384.x.

Wilde, G. J. 2009. *Risk Homeostasis: Theory, Evidence and Practical Implications*. Invited lecture at the Road Safet Advance Course: Department of Civil Engineering, University of Palermo, Italy.

Wohl, J. G. 1961. "Man-Machine Steering Dynamics." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 3: 222–228. doi:10.1177/001872086100300402.

Zhai, S., J. Accot, and R. Woltjer. 2004. "Human Action Laws in Electronic Virtual Worlds: An Empirical Study of Path Steering Performance in VR." *Presence: Teleoperators and Virtual Environments* 13: 113–127. doi:10.1162/1054746041382393.

Appendix – Correlation matrices

Table A1. *Spearman rank-order correlation matrix for the self-paced condition (SP).*

Dependent measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Mean speed	1.00																	
2. Percentage time off-road	0.24	1.00																
3. Mean absolute lateral error	0.10	0.75	1.00															
4. Maximum absolute lateral error	0.24	0.48	0.60	1.00														
5. Self-reported task effort (SRTE)	-0.08	-0.22	-0.32	-0.23	1.00													
6. Median TLC	-0.50	-0.73	-0.45	-0.41	-0.10	1.00												
7. TLC15th	-0.63	-0.78	-0.51	-0.43	-0.08	0.91	1.00											
8. Mean GSR	-0.01	-0.31	-0.19	-0.15	-0.15	0.37	0.32	1.00										
9. Mean GSR rate	0.14	-0.01	-0.06	0.13	0.07	-0.04	-0.11	0.74	1.00									
10. Mean HR	-0.01	-0.31	-0.40	-0.11	0.25	0.34	0.17	0.10	0.10	1.00								
11. SDNN	0.01	0.28	0.33	0.18	0.04	-0.33	-0.16	-0.14	-0.13	-0.58	1.00							
12. LF/HF ratio	0.08	-0.14	-0.21	0.09	0.25	0.09	-0.08	0.03	0.16	0.44	-0.52	1.00						
13. Steering reversal rate	0.08	0.24	-0.05	-0.06	0.33	-0.64	-0.41	-0.11	0.13	-0.30	0.37	-0.13	1.00					
14. Overall NASA TLX	-0.49	0.18	0.19	-0.02	0.31	0.06	0.07	-0.05	-0.06	0.07	0.05	-0.23	0.11	1.00				
15. DSSQ Engagement	0.14	0.02	0.07	0.11	0.02	-0.24	-0.07	-0.28	-0.16	-0.27	0.05	0.16	0.30	0.00	1.00			
16. DSSQ Distress	-0.46	0.11	0.04	-0.16	0.31	0.15	0.12	-0.07	-0.18	0.30	-0.03	-0.13	-0.09	0.84	-0.26	1.00		
17. DSSQ Worry	-0.52	-0.15	0.04	-0.02	0.27	0.38	0.39	0.05	-0.10	0.41	0.06	-0.11	-0.23	0.51	-0.32	0.66	1.00	
18. DBQ	0.25	0.04	-0.29	-0.50	0.18	-0.11	-0.14	-0.02	-0.06	0.14	-0.08	0.01	0.15	-0.21	-0.05	-0.10	-0.30	1.00
19. Mileage	0.58	0.02	0.04	0.09	0.02	-0.25	-0.27	0.06	0.15	-0.07	0.14	0.09	0.11	-0.35	0.33	-0.32	-0.47	0.24

Note. $p < 0.05$ for $|\rho| \geq 0.41$, $p < 0.01$ for $|\rho| \geq 0.52$ and $p < 0.001$ for $|\rho| \geq 0.63$.

Table A2. *Spearman rank-order correlation matrix for the forced-paced condition at 90 km/h (FP90).*

Dependent measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Mean speed																		
2. Percentage time off-road		1.00																
3. Mean absolute lateral error		0.70	1.00															
4. Maximum absolute lateral error		0.48	0.76	1.00														
5. Self-reported task effort (SRTE)		-0.23	-0.26	-0.35	1.00													
6. Median TLC		-0.75	-0.40	-0.34	-0.18	1.00												
7. TLC15th		-0.81	-0.57	-0.49	-0.05	0.93	1.00											
8. Mean GSR		-0.31	-0.14	-0.09	-0.22	0.44	0.31	1.00										
9. Mean GSR rate		-0.09	-0.11	-0.05	-0.12	0.15	0.01	0.82	1.00									
10. Mean HR		-0.45	-0.20	-0.26	0.15	0.38	0.35	0.27	0.16	1.00								
11. SDNN		0.40	0.01	-0.03	-0.13	-0.27	-0.30	-0.08	-0.03	-0.41	1.00							
12. LF/HF ratio		-0.31	-0.15	-0.19	0.02	0.35	0.31	-0.03	0.03	0.45	-0.41	1.00						
13. Steering reversal rate		0.39	0.07	0.16	0.20	-0.71	-0.66	-0.09	0.16	-0.37	0.36	-0.12	1.00					
14. Overall NASA TLX		0.43	0.27	0.08	0.24	-0.45	-0.44	0.09	0.19	0.28	0.13	0.00	0.25	1.00				
15. DSSQ Engagement		0.23	0.25	0.19	-0.13	-0.07	-0.19	-0.19	-0.11	-0.14	0.14	0.16	0.17	-0.02	1.00			
16. DSSQ Distress		0.04	-0.15	-0.12	0.33	-0.19	-0.05	0.20	0.13	0.33	0.00	-0.18	0.12	0.71	-0.36	1.00		
17. DSSQ Worry		-0.14	-0.06	-0.36	-0.07	0.38	0.32	0.40	0.13	0.26	0.27	0.04	-0.12	0.15	0.01	0.18	1.00	
18. DBQ		-0.05	-0.37	-0.45	0.28	-0.14	-0.08	-0.14	0.07	0.30	-0.14	0.09	0.08	0.07	-0.06	0.07	-0.13	1.00
19. Mileage		-0.14	-0.15	-0.01	0.02	0.22	0.17	-0.05	0.05	-0.03	0.07	0.00	-0.06	-0.17	0.23	-0.26	-0.09	0.24

Note. $p < 0.05$ for $|\rho| \geq 0.41$, $p < 0.01$ for $|\rho| \geq 0.52$ and $p < 0.001$ for $|\rho| \geq 0.63$. For FP90 no speed correlations are shown because the speed was fixed, and thus no correlations are applicable.

Table A3. *Spearman rank-order correlation matrix for the forced-paced condition at 130 km/h (FP130)*

Dependent measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Mean speed																		
2. Percentage time off-road		1.00																
3. Mean absolute lateral error		0.75	1.00															
4. Maximum absolute lateral error		0.64	0.78	1.00														
5. Self-reported task effort (SRTE)		-0.06	-0.18	0.05	1.00													
6. Median TLC		-0.76	-0.59	-0.57	-0.02	1.00												
7. TLC15th		-0.78	-0.67	-0.63	0.05	0.91	1.00											
8. Mean GSR		-0.26	-0.35	-0.28	-0.05	0.32	0.31	1.00										
9. Mean GSR rate		0.08	-0.17	0.01	0.01	-0.05	0.04	0.72	1.00									
10. Mean HR		-0.19	-0.22	-0.15	0.03	0.30	0.21	0.07	-0.06	1.00								
11. SDNN		0.04	-0.01	0.05	-0.05	-0.24	-0.06	0.33	0.30	-0.51	1.00							
12. LF/HF ratio		-0.43	-0.35	-0.23	0.07	0.39	0.37	0.12	0.11	0.62	-0.33	1.00						
13. Steering reversal rate		0.30	0.02	0.15	0.13	-0.73	-0.65	-0.09	0.18	-0.16	0.19	-0.08	1.00					
14. Overall NASA TLX		0.50	0.22	0.46	0.40	-0.48	-0.47	-0.12	0.11	-0.01	-0.17	-0.19	0.34	1.00				
15. DSSQ Engagement		-0.28	-0.16	-0.02	0.08	-0.02	0.02	-0.34	-0.16	0.01	-0.09	0.25	0.32	-0.03	1.00			
16. DSSQ Distress		0.12	-0.16	0.14	0.46	-0.13	-0.03	-0.05	0.13	0.02	-0.23	0.07	0.14	0.61	-0.01	1.00		
17. DSSQ Worry		0.09	0.13	0.32	0.15	0.05	0.05	0.00	-0.05	0.09	-0.06	0.10	0.01	-0.09	-0.14	0.11	1.00	
18. DBQ		-0.06	-0.10	-0.25	-0.07	-0.06	-0.10	0.05	-0.01	0.33	0.01	0.15	0.23	-0.31	-0.17	-0.37	-0.08	1.00
19. Mileage		-0.34	-0.14	-0.19	-0.28	0.26	0.21	0.10	0.02	-0.02	0.08	-0.03	-0.12	-0.48	0.17	-0.51	-0.27	0.24

Note. $p < 0.05$ for $|\rho| \geq 0.41$, $p < 0.01$ for $|\rho| \geq 0.52$ and $p < 0.001$ for $|\rho| \geq 0.63$. For FP130 no speed correlations are shown because the speed was fixed, and thus no correlations are applicable.

Table A4. *The results of the repeated measures ANOVA (p, F), Spearman correlation coefficient (ρ) between Run 1 and 3, and Means (M) per dependent measures with the session order as within-subjects factor.*

Dependent measures	*p*-value	*F*(2,46)	*ρ run 1-3*	*M* Run 1	*M* Run 2	*M* Run 3
Mean speed (km/h)	0.955	0.05		113.9	115.5	114.9
Percentage time off-road (%)	0.279	1.31	0.41	6.24	4.85	5.01
Mean absolute lateral error (m)	0.026	3.98	0.69	0.123	0.112	0.119
Maximum absolute lateral error (m)	0.157	1.93	0.39	0.492	0.433	0.435
Self-reported task effort (0-10)	0.407	0.92	0.33	3.73	4.00	3.76
Median TLC (s)	0.892	0.12	-0.05	2.50	2.56	2.53
TLC15th (s)	0.678	0.39	-0.34	1.25	1.28	1.28
Mean GSR (μS)	0.955	0.05	0.81	7.68	7.48	7.61
Mean GSR rate (μS/min)	0.209	1.62	0.86	8.70	7.37	7.36
Mean HR (bpm)	$5.14 \cdot 10^{-5}$	12.33	0.90	80.62	77.14	76.52
SDNN (ms)	0.851	0.16	0.75	56.61	53.48	56.13
LF/HF ratio	0.905	0.10	0.92	1.15	1.11	1.12
Steering reversal rate (reversals/sec)	0.923	0.08	0.63	0.66	0.61	0.61
Overall NASA TLX (%)	0.849	0.16	0.27	50.80	51.77	50.24
DSSQ Engagement (0-32)	0.711	0.34	0.67	23.54	24.04	24.83
DSSQ Distress (0-32)	0.152	1.96	0.31	12.71	13.04	10.79
DSSQ Worry (0-32)	$5.50 \cdot 10^{-3}$	8.88	0.73	7.88	5.54	5.21