Is Improved Lane Keeping during Cognitive Load Caused by Increased Physical Arousal or Gaze Concentration toward the Road Center?

Penghui Li[a, b, *], Gustav Markkula[b], Yibing Li[a, *], Natasha Merat[b]

[a] State Key Laboratory of Automotive Safety & Energy, Department of Automotive Engineering, Tsinghua University, China

[b] Institute for Transport Studies, University of Leeds, UK

* Corresponding author

E-mail address: liph2013@163.com (Penghui Li)

E-mail address: liyb@mail.tsinghua.edu.cn (Yibing Li)

Abstract

Driver distraction is one of the main causes of motor-vehicle accidents. However, the impact on traffic safety of tasks that impose cognitive (non-visual) distraction remains debated. One particularly intriguing finding is that cognitive load seems to improve lane keeping performance, most often quantified as reduced standard deviation of lateral position (SDLP). The main competing hypotheses, supported by current empirical evidence, suggest that cognitive load improves lane keeping via either increased physical arousal, or higher gaze concentration toward the road center, but views are mixed regarding if, and how, these possible mediators influence lane keeping performance. Hence, a simulator study was conducted, with participants driving on a straight city road section whilst completing a cognitive task at different levels of difficulty. In line with previous studies, cognitive load led to increased physical arousal, higher gaze concentration toward the road center, and higher levels of micro-steering activity, accompanied by improved lane keeping performance. More importantly, during the high cognitive task, both physical arousal and gaze concentration changed earlier in time than micro-steering activity, which in turn changed earlier than lane keeping performance. In addition, our results did not show a significant correlation between gaze concentration and physical arousal on the level of individual task recordings. Based on these findings, various multilevel models for micro-steering activity and lane keeping performance were conducted and compared, and the results suggest that all of the mechanisms proposed by existing hypotheses could be simultaneously involved. In other words, it is suggested that cognitive load leads to: (i) an increase in arousal, causing increased micro-steering activity, which in turn improves lane keeping performance, and (ii) an increase in gaze concentration, causing lane keeping improvement through both (a) further increased micro-steering activity and (b) a tendency to steer toward the gaze target.

Keywords: cognitive distraction; cognitive load; physical arousal; gaze concentration; lane keeping improvement; multilevel regression

1. Introduction

Driving is a highly complex task that requires continual integration of perception, cognition, and motor response (Salvucci and Liu, 2002). However, in recent years, with the extensive application of in-vehicle intelligent systems such as navigation devices and mobile/cell phones, driving is now regularly accompanied by engagement in other competing secondary tasks. For instance, a North American survey conducted in 2013, involving 6016 interviewees, showed that 48% of drivers reported answering their cell phone when driving at least some of the time (Schroeder and Meyers et al., 2013). This induced driver distraction introduces many problems for driving safety (Ranney and Mazzae et al., 2000). According to a report by the National Highway Traffic Safety Administration (National Center for Statistics and Analysis, 2016), in the US, distraction-affected crashes contributed to 10% of fatal crashes, 18% of injury crashes, and 16% of all police-reported motor vehicle traffic crashes in 2014. More seriously, results from the 100-Car Naturalistic Driving Study reported that 78% of crashes, and 65% of near-crashes involved driver inattention, including secondary task distraction, driving-related inattention to the forward roadway, drowsiness, and non-specific eye-glances away from the forward roadway (Klauer and Dingus et al., 2006). Hence, it is of importance to investigate the impact of driver distraction on driving performance and its causation.

Driver distraction is commonly defined as *'the diversion of attention away from activities critical for safe driving toward a competing activity'* (Regan and Lee et al., 2008). Two main components are cognitive and visual distraction, described as "mind off road" and "eyes off road" respectively (Victor, 2005). The effect of these activities on driving has been widely explored in recent years (Lamble and Kauranen et al., 1999; Ranney and Mazzae et al., 2000; Engström and Johansson et al., 2005; Jamson and Merat, 2005; Liang and Lee, 2010; Muhrer and Vollrath, 2011; Kountouriotis and Wilkie et al., 2015; Kountouriotis and Merat, 2016), where the consequences of visual distraction on lateral driving performance, and its causation are relatively clear. That is, compared to baseline conditions, visual distraction degrades lateral control (Angell and Auflick et al., 2006; Liang and Lee, 2010; Kountouriotis and Merat, 2016), leading to a significant increase in the standard deviation of lateral position (SDLP, Liang and Lee, 2010; Kountouriotis and Merat, 2016), higher risk of lane departure (Liang and Lee, 2010), and a reduction in time-to-line crossing (Engström and Johansson et al., 2005). These are considered to be due to increased eyes off-road glances during completion of visually distracting tasks (Liang and Lee, 2010; Kountouriotis and Merat, 2016).

However, the effect of cognitive distraction on driving performance is currently unclear (He and McCarley et al., 2014; Kountouriotis and Merat, 2016). In the experimental/laboratory based studies, this kind of distraction is usually triggered by sound-based, cognitively loading, non-visual tasks. Although studies carried out on driving simulators generally suggest that cognitive load impairs driving performance due to the degeneration in drivers' event detection performance (Patten and Kircher et al., 2006; Reyes and Lee, 2008; Haque and Washington, 2014), Naturalistic Driving Studies (NDS) show a mix of unchanged (Olson and Hanowski et al., 2009) and reduced (Victor and Dozza et al., 2015) crash risk during (hands-free) listening or talking on a mobile phone, as reviewed by Carsten & Merat (2015).

There is relatively consistent agreement across the majority of these studies in terms of lane keeping, showing an improvement in performance during cognitively loading tasks, based on reduced SDLP (Engström and Johansson et al., 2005; Jamson and Merat, 2005; Liang and Lee, 2010; Kaber and Liang et al., 2012; He and McCarley et al., 2014; Kountouriotis and Spyridakos et al., 2016). Cognitive load has also been found to lead to increased micro-steering activity (Boer and Rakauskas et al., 2005; Markkula and Engström, 2006; Kountouriotis and Spyridakos et al., 2016; Li and Merat et al., 2017), higher gaze concentration to the forward road center (Victor and Harbluk et al., 2005; Reimer, 2009; Wang and Reimer et al., 2014), and higher physical arousal (Reimer and Mehler, 2011; Mehler and Reimer et al., 2012).

Different hypotheses have been put forward to explain this set of observations during cognitively loading tasks (He and McCarley et al., 2014). Engström et al. (2017) provide an overview in their review, and discuss which hypotheses remain compatible with the available empirical data. Here, we will only consider those hypotheses which remain unrefuted, as shown in Fig. 1.

Engström et al. (2017) suggest the global arousal hypothesis: that improvement in lane keeping is a byproduct of increased cortical arousal during non-automatized tasks, such as those caused by a cognitively loading task. This increased arousal then allows the driver's highly automatized lane keeping and steering behavior to be more responsive to visual stimuli which help support lane keeping in the driving environment, resulting in more frequent micro-steering corrections, in turn leading to reduced SDLP.

Alternatively, the active gaze hypothesis (Wilkie and Wann et al., 2008), also termed the steer-where-you-look hypothesis (Wilson and Chattington et al., 2008; Medeiros-Ward and Seegmiller et al., 2010; He and McCarley et al., 2014), explains the lane keeping improvement as a side effect of task-induced gaze concentration, combined with drivers' tendency to steer in the direction of their gaze. A related suggestion, the visual enhancement hypothesis (Engström and Johansson et al., 2005; He and McCarley et al., 2014; Boer and Spyridakos et al., 2016) suggests that cognitive load causes gaze concentration toward the road center, supporting a better perception of visual information in the road center, and thus resulting in a performance gain of steering which finally leads to lane keeping improvement. As discussed by Engström et al. (2017), if these gaze-mediated mechanisms are real, they are not likely to be solely responsible for SDLP reductions under cognitive load, since such reductions have been observed both without associated reductions in gaze concentration (He et al., 2014) and in conditions of experimentally controlled gaze direction (Cooper et al, 2013). However, these gaze-mediated mechanisms could still be in play, in combination with other non-gaze-mediated mechanisms, such as global arousal.

Thus, the current understanding in this area is that cognitive load affects lane keeping performance via a mediating factor of either physical arousal, gaze concentration toward the road center, or both, with different predictions made by the three competing hypotheses, as shown in Fig. 1. This study presents the first direct tests of these predictions, to investigate the causal relationship suggested by the three hypotheses.

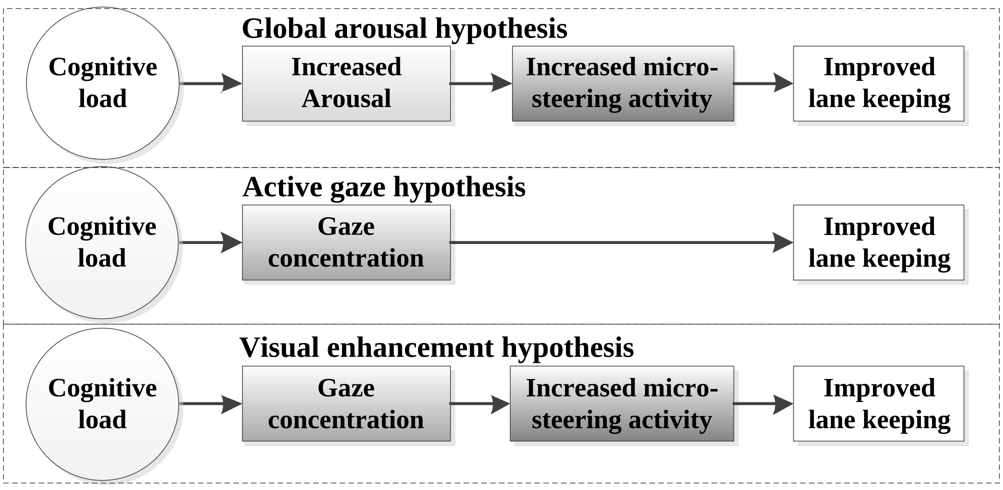

Fig. 1. The main hypotheses used to explain the improved lane keeping performance observed during cognitive load. All boxes are measurable metrics, and the arrows represent predictions. For example, the global arousal hypothesis predicts that increased physical arousal is associated with increased micro-steering activity, which in turn improves lane keeping performance.

In an initial analysis step, a time course method was used to investigate the changes in lane keeping performance, micro-steering activity, gaze concentration, and physical arousal during cognitive task performance. This analysis of change over time provided a first insight into the possible causal relationships between these measures, by means of their temporal patterns of change. Second, pairwise associations between the measures were investigated by univariate multilevel regression, on a sample by sample level, to further constrain the possible causal relationships. Third, a series of multilevel models for micro-steering activity and lane keeping performance, with explanatory variables as proposed by the three competing hypotheses, were conducted and then compared, allowing a final conclusion regarding the possible causal relationships.

2. Method

2.1 Participants

35 participants were recruited using an internet-based forum and by via poster advertisements distributed in Beijing, China. All of them held a valid driving license for a minimum of 2 years, and had normal or corrected-to-normal vision. A within-subjects design was used for the experiment. Our results are based on data from 27 participants (10 females and 17 males), since 3 participants failed to complete the cognitive task experiment because of motion sickness, 3 participants' eye movement data were not adequately recorded, and 2 participants' skin conductance data were also not adequately recorded. The remaining included participants were aged between 20 and 60 years old (mean=35 years, SD=13.5 years).

2.2 Apparatus

The experiment was conducted on a 6 degree-of-freedom motion-based driving simulator, recording data at 60 Hz, in the State Key Laboratory of Automotive Safety and Energy at Tsinghua University, China (see Fig. 2). It is surrounded by 5 screens, providing 200 degrees horizontal and 50 degrees vertical view of the forward road scene, and 36 degrees horizontal and 30 degrees vertical view of the rearward scene, through the rear-view mirror. SensoMotoric Instruments (SMI) eye tracking glasses collected eye movement data, including gaze position, pupil size, and gaze vector, at 30 Hz. A BIOPAC MP150 device was used to record participants' skin conductance level at 100 Hz, at the tips of the left forefinger and mid finger.

Fig. 2. Six degree-of-freedom motion-based driving Simulator

2.3 Secondary tasks

The n-back working memory task, first introduced in a similar driving experiment by researchers at MIT Agelab (Reimer, 2009; Reimer and Gulash et al., 2014), was used as a secondary task. This task requires participants to perform delayed verbal recall of a sequence of digits, which are played to them while driving. In this study, the task was presented at three levels of difficulty: 0-back (low) requires participants to immediately repeat the number they hear, 1-back (medium) requires participants to recall the number one back in the sequence, and 2-back (high) requires participants to recall the number two back in the sequence.

At the start of the task, a message announcing "0 (or 1, 2)-back task begins now" was presented, after which 10 numbers were presented in turn, at a rate of one every 2.25 s, producing a total task length of 34 s.

2.4 Driving environment and experiment design

The driving scenario was a car following task, on a straight section of urban road, which comprised of two motor-lanes, one bicycle-lane and one sidewalk, in each direction. The motor-lanes were 3.5 m wide, with a speed limit of 70 km/h. Several intersections were located on the road, with an interval of 3 km, separating the drive into 4 main experimental blocks. Participants were instructed to follow a lead vehicle, which was driving at a constant speed of 55 km/h, at a comfortable distance, as they usually would during their daily driving. The traffic lights for all intersections remained green at points where the lead vehicle and ego vehicle approached the crossing. The purpose of the light-controlled intersections was to increase the realism of the simulated driving, but the driving data from the intersections were not included in the analyses here.

As outlined above, there were three levels of cognitive task (driving with 0-back, 1-back, and 2-back) and a baseline driving condition. As shown in Fig. 3, in each block (between two intersections), the distraction task was repeated three times, and the interval between every two neighboring tasks was longer than 1 km (30 s), so that participants had enough time to recover from the previous task. One drive, with the four blocks appearing in a random order, lasted about 15 minutes.

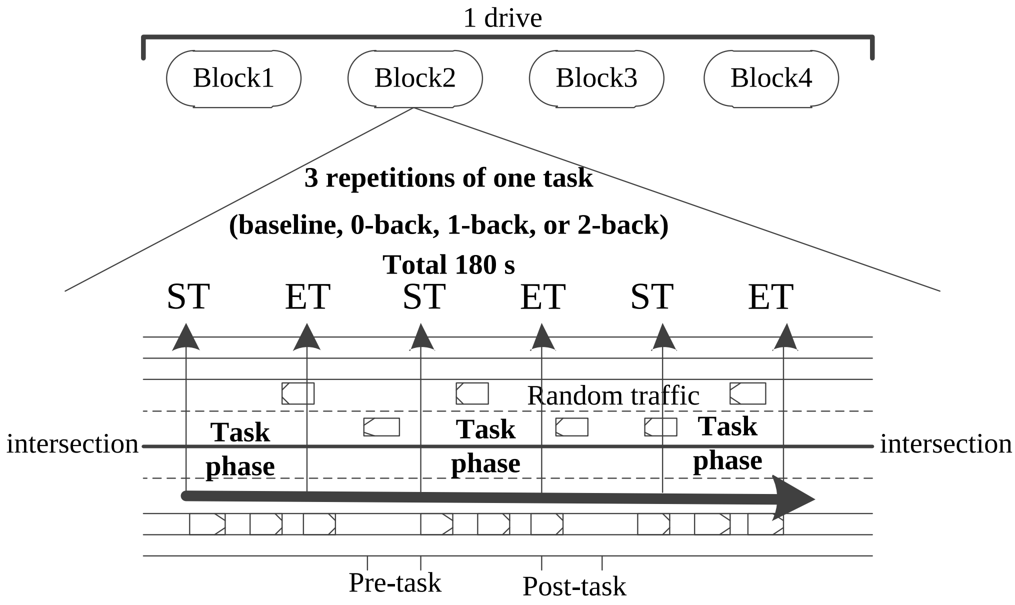

Fig. 3. Experiment design. ST means task start, ET means task end.

2.5 Procedure

The whole experiment consisted of 4 drives, lasting 120 minutes for each participant. Only one drive was accompanied by the cognitive secondary task, which is the main focus of the present study. Half of the participants completed this drive first, and the remainder completed the task in their second drive, providing a counter-balanced study. Each participant's other drives required concurrent completion of a visually distracting task – Arrow test (Engström and Johansson et al., 2005), which will not be reported here.

After arriving at the laboratory, participants were told that their driving behavior would be examined in this experiment and they would complete a training session and four experiment drives. They then received training on the n-back task for 10 minutes, as well as on the other, visual, task. Here, participants were told that their main focus should be on ensuring safe completion of the driving task (as they would in a real driving condition), performing the secondary tasks when they felt safe to do so. Participants were then introduced to the simulator, and provided with about 15 minutes' training. After a short break, participants were equipped with eye tracking glasses and the BIOPAC, and the study commenced. At the end of the experiment, participants completed a questionnaire about their basic personal information, and received 120 RMB for taking part in the study.

2.6 Data analysis

2.6.1 Metrics of driving performance, physical arousal, and gaze concentration

Driving performance, physical arousal, and gaze concentration were measured and analyzed in the present study. The driving performance measures include lane keeping performance and micro-steering activity. Lane keeping performance is most often measured using SDLP, with lower SDLP interpreted as improved lane-keeping performance (He and McCarley et al., 2014).

Micro-steering activity is measured using steering reversal rate (SRR) with a relatively low threshold of 0.5°. An increased SRR0.5° signifies increased micro-steering activity (Markkula and Engström, 2006; Kountouriotis and Spyridakos et al., 2016).

Driver physical arousal is usually measured by skin conductance (Reimer and Mehler, 2011; Braithwaite and Watson et al., 2013). The skin conductance signal can be separated into a background tonic component (Skin Conductance Level: SCL) and a rapid phasic component (Skin Conductance Responses: SCR), both of which result from sympathetic neuronal activity (Braithwaite and Watson et al., 2013). Specifically, SCL relates to the slower acting components and background characteristics of the signal in the absence of SCR, while SCR refers to the faster changing elements of the signal elicited by artifact or stimulus. Both SCL (Measured by the absolute value) and SCR (Measured by the number of SCRs in a given duration) have been widely used to estimate arousal (Roth and Dawson et al., 2012). In this study, SCL data were first resampled with a frequency of 10 Hz, then low-frequency movement artifacts were manually removed (Mehler and Reimer et al., 2012), using the SC analysis software Ledalab (www.ledalab.de).

Here, the mean of a sliding-window standard deviation of skin conductance (MSDSCL) was used to represent driver physical arousal. The MSDSCL metric is obtained by, first, calculating the standard deviation of skin conductance in a sliding 2 s time window moved by 0.1 s steps, thus producing a time series of the standard deviation of skin conductance. Then, MSDSCL is obtained as the mean value of this time series in each task recording (34 s). A higher value of MSDSCL means that the driver was in a state of higher arousal. In this way, this index could capture both changes in the tonic component (SCL) and phasic changes (SCR). By testing the significance of distraction levels (within-subject design: baseline, 0-back, 1-back, and 2-back) on the three arousal metrics with repeated measures general linear model, MSDSCL (η_p^2=0.235) showed a larger effect size than both SCL (η_p^2=0.133) and SCRs (η_p^2=0.207). Therefore, it is more sensitive to cognitive load than both SCL and SCR, and also more continuous in nature. This metric as such will be the subject of more detailed analyses in a later paper (Li, 2017, in preparation).

Driver visual attention toward the road center was measured by standard deviation of horizontal gaze position (SDGAZE), with a lower SDGAZE representing more visual attention concentrated on the road center, also known as increased gaze concentration (Wang and Reimer et al., 2014). For the gaze data, the SMI eye tracker automatically considered low-quality data as blanks, and classified high-quality data into fixations, saccades, and blinks. In the present study, raw gaze data during blanks and blinks, were treated as invalid data. Gaze concentration was calculated based on the remaining valid raw gaze data, when valid data occupied over 50 % of all data in that period, otherwise, gaze concentration was treated as missing data and would not be involved in the following analysis (Reimer, 2009). Finally, out of the 81 recordings per task, there was one excluded recording for baseline (1.2 %), two excluded for 1-back (2.5 %), and one excluded for 2-back (1.2 %).

2.6.2 Time course analysis

A time course method was used to investigate temporal patterns of change in the analyzed measures before, during, and after the cognitive tasks. Similar methods have been used previously for analyzing lane changing maneuvers (Van Winsum and de Waard et al., 1999;

Salvucci and Liu, 2002) and further for investigating the changing pattern of eye movements, and vehicle movements for a lane change (Salvucci and Liu, 2002). Here, only results from the 2-back task are presented, since the demand from the 2-back cognitively loading task was high enough to illustrate significant changes in the four measures. The 0-back and 1-back tasks showed the same trends as the 2-back task, when compared to baseline (See supplementary Fig. S1), although this was weaker, as illustrated also in other such studies (Engström and Johansson et al., 2005; Reimer, 2009; Mehler and Reimer et al., 2012).

Initially, we extracted aggregated data sequences of SDLP, SRR0.5°, MSDSCL, and SDGAZE, including pre-task (from 25 s in advance to start of the cognitive task), task (duration 34 s), and post-task (from the end of cognitive task to 25 s later) data sequences, from the raw data set. Only the second repetition of each task was used, since the pre-task phase of the first task repetition and the post-task phase of the third task repetition involved driving across intersections, which may have influenced drivers' behavior. Then, we computed the four metrics in a sliding 20 s time window, which was moved by 1 s step in the pre-task, task, and post-task data sequences separately. Finally, aggregate graphs of those data were produced, including both the mean and standard error of all drivers' data, in each time window.

2.6.3 Multilevel regression

To determine associations between the analyzed measures, as possible indicators of a causal relationship, we applied both univariate and multivariate multilevel regressions to the measures. Here, crucially, data from all tasks (baseline, 0-back, 1-back, and 2-back) were included, to cover as wide a range as possible of driver states, from low-intensity to high-intensity cognitive load, and the associated effects on other metrics. The data in this study are of a longitudinal format with multi-observations in individual, causing a heterogeneity problem, which makes the classical regression method unsuitable (Hox and Moerbeek et al., 2010; Snijders, 2011; Cohen and Cohen et al., 2013). Hence, regression analyses with a multilevel model (sometimes referred to as a hierarchical model, linear mixed model, or random effects model) with a random intercept was used to investigate relationships between the measures (Huang and Chin et al., 2008; Huang and Abdel-Aty, 2010).

A 2-level multilevel model with random intercept was used. In level 1, X_{ij} represents the explanatory variable of ith driver in the jth driving situation, and the corresponding response variables Y_{ij} is expressed as (Hox and Moerbeek et al., 2010; Snijders, 2011):

$$Y_{ij} = \beta_{0i} + \beta X_{ij} + R_{ij} \qquad (1)$$

In level 2, the within-driver correlation was taken into consideration. That is, data from the same driver i share the same intercept β_{0i}, while data from different drivers have different intercepts. The formulation is:

$$\beta_{0i} = \beta_0 + U_{0i} \qquad (2)$$

Specifically, this model has both fixed and random terms (Huang and Abdel-Aty, 2010). In the fixed term, β_0 is the average intercept, β is the fixed-effect coefficient of covariates X on the response variable. In the random term, U_{0i} is driver-dependent deviation, representing between-driver variation. R_{ij} represents within-driver variation, and is the disturbance term associated with level 1 analysis. The maximum likelihood method was used for model parameter estimation (Hox and Moerbeek et al., 2010), and the analyses were conducted using Matlab software.

To determine the best model for understanding lane keeping improvement, the model comparison method (Victor and Dozza et al., 2015) was used, with Akaike Information Criterion (AIC) as index. The model with lower AIC can be regarded as significantly better (Akaike, 1998), when AIC difference between two models over 2 (Burnham and Anderson, 2004). For quantitative comparison of a set of nested models, the likelihood-ratio test was also used (Wilks, 1938), with the full model (larger model) preferable to the reduced model (simpler model) if the test reaches $p < 0.05$ significance-level.

3. Results

3.1 Time course analysis of driving performance, physical arousal and gaze concentration

Fig. 4 shows driver physical arousal, gaze concentration, micro-steering activity, and lane keeping performance as a function of time. MSDSCL and SDGAZE both showed notable changes (44.5% and -39.2%, respectively) from the very start of the 2-back task, and these levels remained relatively constant throughout task engagement. SRR0.5° showed a smaller increase (10.4%) at the beginning of the task, and then went on increasing until the end. SDLP had a minor decrease (-3.7%) at task start, then remained relatively constant for a few seconds, before showing a more notable decreasing trend until the end of the task. After the task, all values recovered to their initial, pre-task levels.

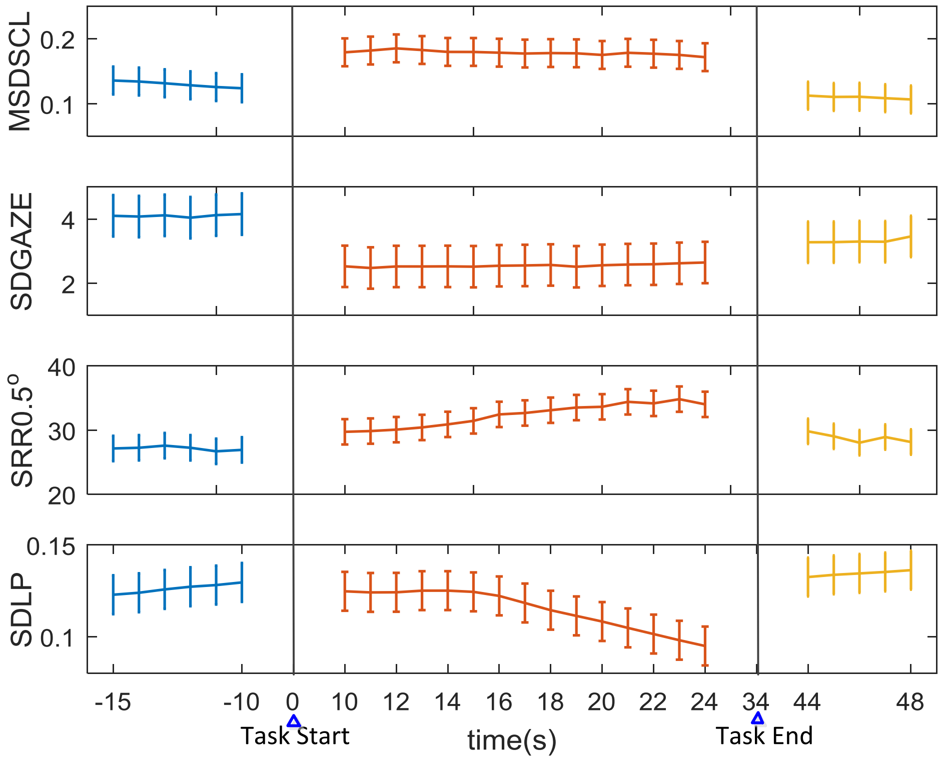

Fig. 4. Generalized time course of physical arousal (mean of standard deviation of skin conductance level, MSDSCL), gaze concentration (standard deviation of gaze yaw angle, SDGAZE), micro-steering activity (steering reversal rate at 0.5° level, SRR0.5°), and lane keeping performance (standard deviation of lane position, SDLP; mean and SE shown in all panels) before, during, and after the 2-back task.

Paired t-tests with 0.05 significance level were used to determine when the measures started showing a significant change during the task phase, compared to the pre-task phase. Since none of the four measures showed significant differences between the 6 time steps of the pre-task phase (all p>0.05), the last data points of the pre-task phase (at -10 s in Fig. 4) were used to represent data in the pre-task phase, and these were compared with the data in the ensuing 15 time steps in the task phase, as shown in Fig. 5. The results illustrate that, both MSDSCL and SDGAZE showed a significant change from the start of the task phase, while a significant SRR0.5° change was observed 5 s after the start, and SDLP 12 s after task start. (Corrections for multiple comparisons have not been made here, since the main point is to illustrate the general order of changes, not any exact timings.)

It can be noted that all three hypotheses mentioned in the Introduction suggest the following sequence of causation: 1) cognitive load, 2) increased arousal or gaze concentration, 3) possibly increased micro steering, and 4) improved lane keeping. In other words, the order of effects observed here is compatible for all three hypotheses. (If, for example, micro steering would have been visible before any changes to arousal or gaze concentration, this would have been problematic for the global arousal or visual enhancement hypotheses, respectively.)

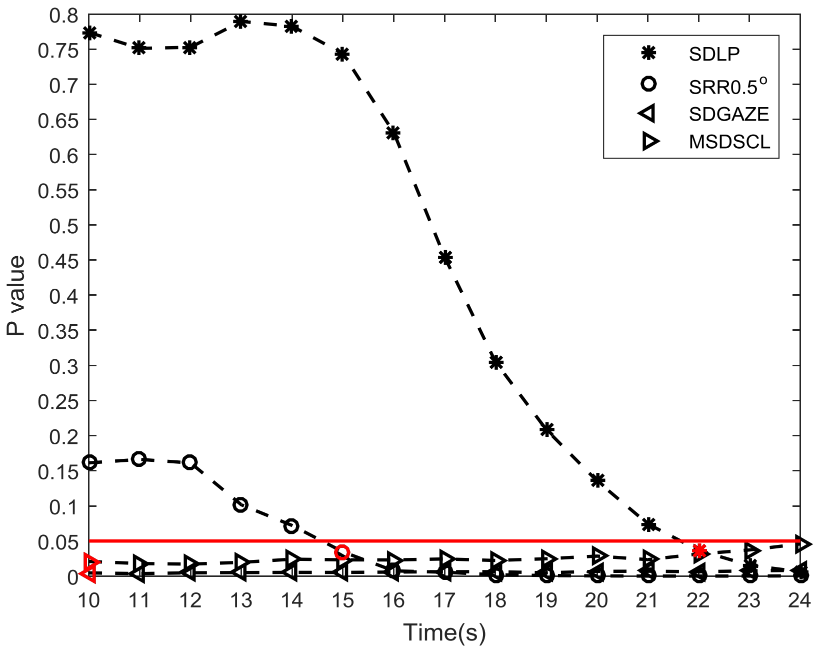

Fig. 5. Time course of p-values for paired t-tests during the 2-back task phase, compared to the pre-task phase (Time in x axis is the same as that in Fig. 4).

It should be mentioned here that the choice of a 20 s time window of analysis moving in 1 s steps was made after we tried a series of time windows, starting from 30 s and narrowing until 20 s. A 30 s (or more) window has been commonly used for these metrics in previous studies (Engström and Johansson et al., 2005; Kountouriotis and Spyridakos et al., 2016). We obtained similar results as previous studies when using 30 s time window. However, we found that this window was too wide to capture the temporal changes in the four variables. At shorter time windows than 20 s, the patterns of temporal change remain the same as in Figs. 4 and 5, except for the SDLP and SRR0.5^{o} metric for very short time windows (less than 10 s); this metric does not work well for time windows shorter than a typical period of vehicle movement in the lane (Östlund and Peters et al., 2005; Li and Merat et al., 2017).

3.2 Pairwise association of driving performance measures and mediators

To further study the possible causal relationships between driving performance measures and the hypothesized mediators, pairwise associations were conducted among SDLP, SRR0.5^{o}, MSDSCL, and SDGAZE. Table 1 shows the descriptive statistics of variables in our data set. Here, we only considered the task phase, and not the pre and post phases, such that we could now include all three task repetitions per participant. One task phase produced one data point for each variable, and across the four driving conditions (baseline and three task difficulties) each participant thus produced 12 data points, for a total of 324 data points from the 27 participants in this study. These were all included in the univariate multilevel regression.

Table 1.
Descriptive statistics of measures of task-driving data set for ML model

Measures	Description	Min	Max	Average	SD	Count
SDLP(m)	Standard deviation of vehicle's lateral position	0.028	0.525	0.150	0.077	324
SRR0.5^{o}(/min)	Steering reversal rate (level 0.5^{o})	0	86	23	12	324
MSDSCL(μS)	Mean of standard deviation of skin conductance level	0.004	0.748	0.164	0.136	324
SDGAZE(^{o})	Standard deviation of horizontal gaze position	0.400	14.663	3.836	3.005	324

Based on the hypotheses investigated here (tentatively supported by the time course analysis), in the following pairwise association analyses, micro-steering activity should be the response variable for physical arousal and gaze concentration respectively. Similarly, lane keeping performance should be the response variable for physical arousal, gaze concentration, and micro-steering activity respectively.

Table 2 shows results with slope and significance level of each univariate multilevel regression. The results showed that, improved lane keeping performance (reduced SDLP) was significantly associated with increased micro-steering activity (increased SRR 0.5^{o}), increased physical arousal (increased MSDSCL), and increased gaze concentration (reduced SDGAZE). Meanwhile, increased micro-steering activity was significantly associated with increased physical arousal and gaze concentration. However, gaze concentration did not show a significant association with physical arousal. These results provide further support for the predictions provided for the three hypotheses. Perhaps most importantly, the lack of a significant association between physical arousal and gaze concentration suggests that the well-documented effects of cognitive tasks on these measures are two at least partially separate phenomena, since, on a per-observation basis, physical arousal and gaze concentration varied independently of each other in our data.

Table 2.

Results of univariate multilevel regression in the presence of cognitive task (Significance level: ** p<0.001, * p<0.05)

Slope & Significance		Independent Variables		
		SRR0.5°	SDGAZE	MSDSCL
Response Variables	SDLP	-0.0015**	0.0045**	-0.1002**
	SRR0.5°		-0.6527**	19.697**
	SDGAZE			-1.946
	MSDSCL		-0.0034	

3.3 Application of multiple multilevel regression to the driving measures

To provide a further, more direct contrast between the three hypotheses, a series of multilevel models for micro-steering activity and lane keeping performance were constructed, based on the hypotheses, and compared.

3.3.1 Micro-steering activity

According to the global arousal hypothesis, increased micro-steering activity during cognitively loading tasks is caused by an enhanced physical arousal, whereas the visual enhancement hypothesis suggests that this rise in micro-steering activity is due to an increase in gaze concentration towards the road center. Our results thus far, from the temporal analyses and the pairwise regressions, are compatible with both hypotheses, but have suggested that there is no causal link between the two involved mediators, physical arousal and gaze concentration. To test whether both mechanisms coexist or not, we therefore constructed three multivariate models of micro-steering activity: (1) a univariate arousal-based model, suggested by the global arousal hypothesis, with only physical arousal as the explanatory variable, (2) a univariate gaze-based model, suggested by the visual enhancement hypothesis, with only gaze concentration as the explanatory variable, and (3) a bivariate arousal-gaze model, based on the hypothesis that both mechanisms are simultaneously active.

Table 3.
Multilevel model of SRR0.5°

Model 1: Arousal-based model (Global arousal hypothesis)									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	19.2830	2.0448	321	9.43	0.0000	9.6 (7.3, 12.7)	6.8 (6.2, 7.3)	2248	0.66
MSDSCL	19.6970	4.7287	321	4.17	0.0000				
Model 2: Gaze-based model (Visual enhancement hypothesis)									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	25.0140	2.0565	321	12.16	0.0000	9.7 (7.3, 12.8)	6.8 (6.3, 7.4)	2255	0.65
SDGAZE	-0.6527	0.2040	321	-3.20	0.0015				
Model 3: Arousal-gaze model (Global arousal hypothesis and Visual enhancement hypothesis)									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	21.7240	2.2781	320	9.54	0.0000	10.1 (7.6, 13.4)	6.6 (6.1, 7.2)	2241	0.67

SDGAZE	-0.6003	0.1996	320	-3.01	0.0028
MSDSCL	18.7920	4.6790	320	4.02	0.0001

Likelihood-ratio test: P(Model-3 vs. Model-1)=0.003; P(Model-3 vs. Model-2)<0.001

As shown in Table 3, Model 1 and 2 are in fact the same as in the pairwise association tests in the Section 3.2. Notably, however, the analyses showed that Model 3 was preferable over model 1 and, both in terms of having the lowest AIC among the three models, and the likelihood-ratio test also showing this model to be significantly better than the other two. Critically, these results suggest two separate causation pathways between cognitive load and micro-steering activity, one involving arousal but not gaze concentration, and one involving gaze concentration but not arousal. Thus, both the global arousal and visual enhancement hypotheses could independently contribute to explaining the variability of micro-steering activity.

3.3.2 Lane keeping performance

As for the cause of reduction in SDLP, both the global arousal and visual enhancement hypotheses suggest that it is caused by increased micro-steering activity, while the active gaze hypothesis suggests a direct causal link from increased gaze concentration tom reduced SDLP. In Section 3.2, associations were provided between SDLP and both SRR 0.5^{o} and SDGAZE, as illustrated by the univariate multilevel regressions. Here, we follow up with multiple multilevel regressions, considering all three hypotheses.

Four main multilevel models of SDLP were constructed, as shown in Table 4: (1) A univariate steering-based model, suggested by both the global arousal and visual enhancement hypotheses, with only micro-steering activity as explanatory variable. (2) A univariate gaze- based model, suggested by the active gaze hypothesis, with only gaze concentration as the explanatory variable. (3) A bivariate steering-gaze model, suggested by the possibility of all three causal pathways being simultaneously active, with both micro-steering activity and gaze concentration as the explanatory variables. To test whether physical arousal contributes to explaining the variability of SDLP due to some other unknown mechanisms, we also tested (4) a trivariate steering-gaze-arousal SDLP model, with micro-steering activity, physical arousal, and gaze concentration as explanatory variables.

Table 4 shows the results for these multilevel models of SDLP. Comparing the four nested multilevel models, AIC values and likelihood ratio tests all indicate that Model 3 is preferable over Models 1 and 2. However, Model 4 was not preferable over Model 3, neither based on AIC (Model 3: -868, Model 4: -869) or the likelihood-ratio test (p=0.051). This suggests that Model 3, the steering-gaze model, is preferable for explaining the variability of SDLP, i.e., that both increased micro-steering activity and gaze concentration contributed to the reduction in SDLP, but without a direct link from arousal to reduction in SDLP.

Table 4.
Multilevel model of SDLP

Model 1: Steering-based model (Global arousal hypothesis and Visual enhancement hypothesis)

Variables	Fixed Effects					Random Effects		AIC	R^2
	Coefficient	SE	DF	tStat	P_value	Intercept variance	Residual		

(Intercept)	0.1824	0.0136	320	13.42	0.0000	0.05 (0.03, 0.06)	0.06 (0.05, 0.06)	-865	0.41
SRR0.5 °	-0.0015	0.0004	320	-3.51	0.0005				
Model 2: Gaze-based model (Active gaze hypothesis)									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	0.1314	0.0120	321	10.93	0.0000	0.05 (0.04, 0.07)	0.06 (0.05, 0.06)	-863	0.41
SDGAZE	0.0045	0.0016	321	2.80	0.0055				
Model 3: Steering-gaze model (Global arousal hypothesis, Visual enhancement hypothesis and Active gaze hypothesis)									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	0.1662	0.0158	319	10.55	0.0000	0.05 (0.04, 0.07)	0.06 (0.05, 0.06)	-868	0.42
SRR0.5 °	-0.0014	0.0004	319	-3.22	0.0014				
SDGAZE	0.0036	0.0016	319	2.20	0.0285				
Model 4: Steering-gaze-arousal model									
Variables	**Fixed Effects**					**Random Effects**		**AIC**	**R^2**
	Coefficient	**SE**	**DF**	**tStat**	**P_value**	**Intercept variance**	**Residual**		
(Intercept)	0.1755	0.0164	318	10.68	0.0000	0.05 (0.04, 0.07)	0.06 (0.05, 0.06)	-869	0.43
SRR0.5 °	-0.0012	0.0004	318	-2.82	0.0050				
SDGAZE	0.0034	0.0016	318	2.10	0.0368				
MSDSCL	-0.0740	0.0377	318	-1.96	0.0507				

Likelihood-ratio test: P(Model-3 vs. Model-1)=0.031; P(Model-3 vs. Model-2)=0.009; P(Model-4 vs. Model-3)=0.051; P(Model-4 vs. Model-1)=0.014; P(Model-4 vs. Model-2)=0.005

4. Discussion

The aim of the present study was to further clarify why drivers' lane keeping performance is improved by a concurrent cognitively loading task. Various authors have suggested that cognitive load affects lane keeping performance via physical arousal or gaze concentration as mediators, but hold mixed ideas about whether, and how, those mediators influence lane keeping performance. Here, three difficulty levels of a cognitive task were presented to drivers in a simulator study, and the hypothesized relationships between driving performance measures (lane keeping performance, micro-steering activity) and the possible mediators (physical arousal and gaze concentration) were analyzed.

In line with previous studies, during performance of the cognitive tasks, we observed improved lane keeping performance (Engström and Johansson et al., 2005; Jamson and Merat, 2005; He and McCarley et al., 2014; Kountouriotis and Merat, 2016), increased micro-steering activity (Boer and Rakauskas et al., 2005; Kountouriotis and Spyridakos et al., 2016), increased gaze concentration to the forward road center (Victor and Harbluk et al., 2005; Reimer, 2009; Wang and Reimer et al., 2014), and increased physical arousal (Reimer and Mehler, 2011; Mehler and Reimer et al., 2012).

In addition, a time course analysis, of the most demanding version of the cognitive task, showed, for the first time, the temporal dynamics of lane keeping performance, micro-steering activity, gaze concentration, and physical arousal. That is, before engaging in a cognitive task, driver lane keeping performance was more erratic, and micro-steering activity, gaze concentration toward the road center, and physical arousal were at a relatively low level. The start of the cognitive task caused an immediate increase in gaze concentration and physical arousal, but the change in micro-steering activity and lane keeping performance was more gradual. Results of paired t-tests showed that, during a high cognitive task, gaze concentration and physical arousal changed earlier than micro-steering activity, which in turn changed earlier than lane keeping performance. This implies that the effect of cognitive load on gaze concentration and physical arousal precedes that of micro-steering activity which in turn affects lane keeping performance. This aligns nicely with the idea of gaze concentration and/or physical arousal being the cause of increased micro-steering activity, and some or all of gaze concentration, physical arousal, and micro-steering activity being the cause of lane keeping improvement, exactly as proposed by the three hypotheses investigated here: the global arousal, visual enhancement, and active gaze hypotheses, thus providing support for constraining our subsequent analyses to the specific causal relationships proposed by these hypotheses.

Using univariate multilevel regression analyses, we then demonstrated the existence of all the pairwise associations between measures predicted by the three hypotheses. However, we did not find any association between the two mediating variables physical arousal and gaze concentration. This is interesting given that the classical cue utilization hypothesis (Easterbrook, 1959) suggests that increased arousal reduces the range of cue utilization, which has led some authors to suggest that gaze concentration during cognitive load is caused by increased arousal (Monk and Yang et al., 2013). We did not find any evidence for such an association, and therefore suggest that arousal and gaze concentration are, at least to some extent, independently affected by cognitive load. Further research on the relationship between gaze concentration and arousal seems warranted.

Based on these findings, and further supported by additional multilevel models, two independent pathways of cognitive load affecting lane keeping performance were identified here, as illustrated in Fig. 6: (1) An increase in arousal, causing increased micro-steering activity, which in turn improves lane keeping performance, as suggested by the global arousal hypothesis. (2) Gaze concentration, causing improved lane keeping performance both through (a) increased micro-steering activity, as suggested by the visual enhancement hypothesis, and (b) a tendency to steer toward the gaze target, as suggested by the active gaze hypothesis.

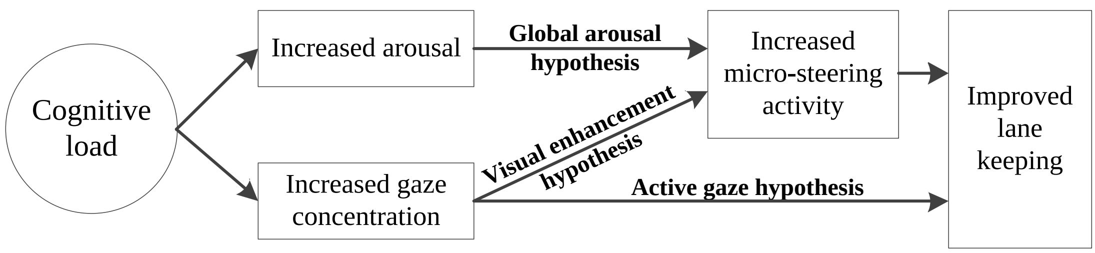

Fig. 6. Structure of causation of lane keeping improvement during cognitive load

To further test this conclusion, some additional tests were carried out. First, if the causation structure proposed in Fig. 6 is correct, our best multivariate model, the steering-gaze model for SDLP, should perform at least as well or better as a model predicting SDLP from cognitive task difficulty. Therefore, we also constructed such a task-level based model, with 0-back, 1-back, and 2-back as nominal variables (the same as the well-known repeated measures general linear model), and baseline driving as a constant variable (model intercept). As shown in Table 5, this model produces a similar but slightly higher (worse) AIC than the steering-gaze model (-867 and -868, respectively). This difference in AIC is not considered to be statistically significant (Burnham and Anderson, 2004). In other words, the steering-gaze mixed model does not explain the variability of SDLP significantly better than the task-level-based model, but it does explain it in a way that sheds more light on the involved mechanisms and mediators.

Table 5.
Task-level-based model of SDLP

Variables	Fixed Effects					Random Effects		AIC	R^2
	Coefficient	SE	DF	tStat	P_value	Intercept variance	Residual		
(Intercept)	0.1674	0.0111	319	15.05	0.0000	0.05 (0.04, 0.06)	0.06 (0.05, 0.06)	-867	0.42
0back	-0.0173	0.0089	319	-1.94	0.0527				
1back	-0.0227	0.0089	319	-2.55	0.0113				
2back	-0.0349	0.0089	319	-3.91	0.0001				

Second, we also applied Structural Equation Modelling (SEM) to further test the proposed structure in Fig. 6. The reader may note that in practice, the comparison of multiple multivariate regressions is in fact very similar to an SEM analysis. The reason we did not perform SEM from the outset is that, we were not able to find any existing SEM software that would permit our combination of discrete variables (cognitive load) and continuous variables (the various metrics). Therefore, the SEM included the full structure in Fig. 6, except the cognitive load variable, i.e., it started from gaze concentration and arousal (in practice, since in our case all variables were directly measured, this was a special case of SEM referred to as Path Analysis). The fit indices of this SEM model showed that it was acceptable, supporting the proposed structure for understanding lane keeping improvement in the present study.

In summary, this study has provided evidence suggesting that the phenomenon of reduced lane position variability during cognitively loading tasks is highly complex and multifaceted, with all of the mechanisms shown in Fig. 6 being simultaneously active. In other words, our results suggest that all three of the main hypotheses considered in this paper are true.

However, it should be emphasized that our time course and regression analyses do not provide conclusive proof of causation. The reason we have nevertheless discussed our findings in causal terms is because we departed from three well-defined causal hypotheses, which could each in theory have been refuted by our data and analyses. What we have shown in practice is that our data were compatible with the idea of all three tested causal hypotheses being true at the same time. This does not, however, preclude the possible existence of other mechanisms or hidden mediating factors. To further test whether the proposed causal structure in Fig. 6 is sufficient, one path forward would be to conduct more targeted experiments; e.g., manipulating arousal and gaze concentration directly, to see if this, in itself, leads to improvements in lane keeping. In such studies one could also consider investigating the time course of the involved behaviors in even more detail than what has been done here. For example, the difference in time of onset between micro-steering increases and improved lane keeping observed here (Fig. 5) is relatively large. Further research is warranted to establish if this is because it takes a long time for these small steering actions to translate into an actual impact on the vehicle's path, or because the putative parallel effect of gaze concentration on the lane keeping builds up relatively slowly over time.

Also, with respect to the time course analysis, even though the overall pattern in 0- and 1-back tasks were similar to those in the 2-back task (i.e., the overall effects in the task phase are pointing in same direction- but with smaller effect sizes - See supplementary Fig. S1), there were indications of possible time course patterns that were not present in the 2-back data, such as MSDSCL falling over time during 0-back, and SRR0.5^{o} seemingly not increasing over time in the 0- and 1-back task (but still with indications of a small increase from the start of the task, as in the 2-back data).

From a more applied perspective, our finding that physical arousal and gaze concentration change earlier than driving performance measures (micro-steering activity, and lane keeping performance) during a cognitive task, suggests that both physical arousal and gaze concentration could be used for early detection of cognitive load effects on driving, before driving measures have started to change. The finding from (Reimer and Mehler et al., 2011) that physiological measures such as physical arousal and gaze concentration are more sensitive to changes in driver workload than driving performance measure also aligns with this idea.

Acknowledgements

This work was supported by the Independent Project of State Key Laboratory of Automotive Safety & Energy (ZZ2016-032), Tsinghua University, China. In addition, Penghui Li was funded by both Tsinghua Scholarship for Overseas Graduate Studies (2016049), and Leeds Researcher Mobility Award.

References

Akaike, H., 1998. Information theory and an extension of the maximum likelihood principle, Springer: 199-213.

Angell, L. S., Auflick, J., Austria, P. A., Kochhar, D. S., Tijerina, L., Biever, W.,... Kiger, S., 2006. Driver workload metrics task 2 final report. No. HS-810 635. 2006.

Boer, E. R., Rakauskas, M. E., Ward, N. J., & Goodrich, M. A., 2005. Steering entropy revisited. Proceedings of the 3rd International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Boer, E. R., Spyridakos, P. D., Markkula, G. M., & Merat, N., 2016. "Cognitive Driver Distraction Improves Straight Lane Keeping: A Cybernetic Control Theoretic Explanation." IFAC-PapersOnLine **49** (19): 627-632.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M., 2013. "A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments." Psychophysiology **49**: 1017-1034.

B Burnham, K. P., & Anderson, D. R., 2004. "Multimodel inference: understanding AIC and BIC in model selection." Sociological methods & research **33** (2): 261-304.

Carsten, O., & Merat, N., 2015. "Protective or Not?" 4th International Driver Distraction and Inattention Conference, Sydney, New South Walesfrom.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S., 2013. Applied multiple regression/correlation analysis for the behavioral sciences, Routledge.

Cooper, J.M., Medeiros-Ward, N. and Strayer, D.L., 2013. "The Impact of Eye Movements and Cognitive Workload on Lateral Position Variability in Driving." Human Factors: The Journal of the Human Factors and Ergonomics Society **55** (5): 1001 -1014.

Easterbrook, J. A., 1959. "The effect of emotion on cue utilization and the organization of behavior." Psychological review **66** (3): 183.

Engström, J., Markkula, G., Victor, T., Merat, N., 2017. "Effects of cognitive load on driving performance: The cognitive control hypothesis." Human Factors **59** (5): 734-764.

Engström, J., Johansson, E., & Östlund, J., 2005. "Effects of visual and cognitive load in real and simulated motorway driving." Transportation Research Part F: Traffic Psychology and Behaviour **8** (2): 97-120.

Haque, M. M., & Washington, S., 2014. "A parametric duration model of the reaction times of drivers distracted by mobile phone conversations." Accident Analysis & Prevention **62**: 42-53.

He, J., McCarley, J. S., & Kramer, A. F., 2014. "Lane Keeping Under Cognitive Load: Performance Changes and Mechanisms." Human Factors: The Journal of the Human Factors and Ergonomics Society **56** (2): 414 -426.

Hox, J. J., Moerbeek, M., & van de Schoot, R., 2010. Multilevel analysis: Techniques and applications, Routledge.

Huang, H., Chin, H. C., & Haque, M. M., 2008. "Severity of driver injury and vehicle damage in traffic crashes at intersections: a Bayesian hierarchical analysis." Accident Analysis & Prevention **40** (1): 45-54.

Huang, H., & Abdel-Aty, M., 2010. "Multilevel data and Bayesian analysis in traffic safety." Accident Analysis & Prevention **42** (6): 1556-1565.

Jamson, A. H., & Merat, N., 2005. "Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving." Transportation Research Part F: Traffic Psychology and Behaviour **8** (2): 79-96.

Kaber, D. B., Liang, Y., Zhang, Y., Rogers, M. L., & Gangakhedkar, S., 2012. "Driver performance effects of simultaneous visual and cognitive distraction and adaptation behavior." Transportation research part F: traffic psychology and behaviour **15** (5): 491-501.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J., 2006. "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data.".

Kountouriotis, G. K. and Merat N., 2016. "Leading to distraction: Driver distraction, lead car, and road environment." Accident Analysis & Prevention **89**: 22-30.

Kountouriotis, G. K., Spyridakos, P., Carsten, O. M., & Merat, N., 2016. "Identifying cognitive distraction using steering wheel reversal rates." Accident Analysis & Prevention **96**: 39-45.

Kountouriotis, G. K., Wilkie, R. M., Gardner, P. H., & Merat, N., 2015. "Looking and thinking when driving: The impact of gaze and cognitive load on steering." Transportation research part F: traffic psychology and behaviour **34**: 108-121.

Lamble, D., Kauranen, T., Laakso, M., & Summala, H., 1999. "Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving." Accident Analysis & Prevention **31** (6): 617-23.

Li, P. (2017). "Characterizing driver's physical arousal using skin conductance level (paper preparation).".

Li, P., Merat, N., Zheng, Z., Markkula, G., Li, Y.,... Wang, Y., 2017. "Does cognitive distraction improve or degrade lane keeping performance? Analysis of time-to-line crossing safety margins." Transportation Research Part F: Traffic Psychology and Behaviour.

Liang, Y., & Lee, J. D., 2010. "Combining cognitive and visual distraction: Less than the sum of its parts." Accident Analysis & Prevention **42** (3): 881-890.

arkkula, G., & Engström, J., 2006. A steering wheel reversal rate metric for assessing effects of visual and cognitive secondary task load. Proceedings of the 13th ITS world congress, London.

Medeiros-Ward, N., Seegmiller, J., Cooper, J., & Strayer, D., 2010. Dissociating eye movements and workload on lateral lane position variability, SAGE Publications.

Mehler, B., Reimer, B., & Coughlin, J. F., 2012. "Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task an on-road study across three age groups." Human Factors: The Journal of the Human Factors and Ergonomics Society **54** (3): 396-412.

Monk, C., Yang, D., McGehee, D., Hanowski, R., Horrey, B., Young, D.,... Regan, M., 2013. Report from the US-EU Focus Group on Cognitive Load. Gothenburg.

Muhrer, E., & Vollrath, M., 2011. "The effect of visual and cognitive distraction on driver's anticipation in a simulated car following scenario." Transportation research part F: traffic psychology and behaviour **14** (6): 555-566.

National Center for Statistics and Analysis, 2016. "Distracted Driving 2014." in Traffic Safety Research Notes.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. L., 2009. Driver distraction in commercial vehicle operations. U.S. Department of Transportation, Washington, D.C, Federal Motor Carrier Safety Administration.

Östlund, J., Peters, B., Thorslund, B., Engström, J., Markkula, G., Keinath, A.,... Foehl, U., 2005. "Driving performance assessment-methods and metrics.".

Patten, C. J., Kircher, A., Östlund, J., Nilsson, L., & Svenson, O., 2006. "Driver experience and cognitive workload in different traffic environments." Accident Analysis & Prevention **38** (5): 887-894.

Ranney, T. A., Mazzae, E., Garrott, R., & Goodman, M. J., 2000. NHTSA driver distraction research: Past, present, and future. Driver distraction internet forum.

Regan, M. A., Lee, J. D., & Young, K., 2008. Driver distraction: Theory, effects, and mitigation, CRC Press.

Reimer, B., 2009. "Impact of cognitive task complexity on drivers' visual tunneling." Transportation Research Record: Journal of the Transportation Research Board(2138): 13-19.

Reimer, B., Mehler, B., Coughlin, J. F., Roy, N., & Dusek, J. A., 2011. "The impact of a naturalistic hands-free cellular phone task on heart rate and simulated driving performance in two age groups." Transportation Research Part F: Traffic Psychology and Behaviour **14** (1): 13-25.

Reimer, B., & Mehler, B., 2011. "The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation." Ergonomics **54** (10): 932-42.

Reimer, B., Gulash, C., Mehler, B., Foley, J. P., Arredondo, S.,... Waldmann, A., 2014. The MIT AgeLab n-back: a multi-modal android application implementation. Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, ACM.

Reyes, M. L., & Lee, J. D., 2008. "Effects of cognitive load presence and duration on driver eye movements and event detection performance." Transportation research part F: traffic psychology and behaviour **11** (6): 391-402.

Roth, W. T., Dawson, M. E., & Filion, D. L., 2012. "Publication recommendations for electrodermal measurements." Psychophysiology **49**: 1017-1034.

Salvucci, D. D., & Liu, A., 2002. "The time course of a lane change: Driver control and eye-movement

behavior." Transportation Research Part F: Traffic Psychology and Behaviour **5** (2): 123-132.
Schroeder, P., Meyers, M., & Kostyniuk, L., 2013. National survey on distracted driving attitudes and behaviors--2012. No. DOT HS 811 729.
Snijders, T. A., 2011. Multilevel analysis, Springer.
Van Winsum, W., de Waard, D., & Brookhuis, K. A., 1999. "Lane change manoeuvres and safety margins." Transportation Research Part F: Traffic Psychology and Behaviour **2** (3): 139-149.
Victor, T. W., Harbluk, J. L., & Engstrom, J. A., 2005. "Sensitivity of eye-movement measures to in-vehicle task difficulty." Transportation research part F: traffic psychology and behaviour **8** (2): 167-190.
Victor, T., 2005. Keeping eye and mind on the road, Acta Universitatis Upsaliensis. **Doctoral dissertation**.
Victor, T., Dozza, M., Bärgman, J., Boda, C., Engström, J., Flannagan, C.,... Markkula, G., 2015. Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. Washington, D.C, Transportation Research Board.
Wang, Y., Reimer, B., Dobres, J., & Mehler, B., 2014. "The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand." Transportation Research Part F: Traffic Psychology and Behaviour **26, Part A**: 227-237.
Wilkie, R. M., Wann, J. P., & Allison, R. S., 2008. "Active gaze, visual look-ahead, and locomotor control." Journal of experimental psychology: Human perception and performance **34** (5): 1150.
Wilks, S. S., 1938. "The large-sample distribution of the likelihood ratio for testing composite hypotheses." The Annals of Mathematical Statistics **9** (1): 60-62.
Wilson, M., Chattington, M., & Marple-Horvat, D. E., 2008. "Eye movements drive steering: Reduced eye movement distribution impairs steering and driving performance." Journal of motor behavior **40** (3): 190-202.