



# Event-related potentials as indices of mental workload while using an in-vehicle information system

I. Solís-Marcos<sup>1</sup> · K. Kircher<sup>1</sup>

Received: 11 September 2017 / Accepted: 18 April 2018 / Published online: 24 April 2018  
© The Author(s) 2018

## Abstract

New in-vehicle information systems are now being commercialized. Despite the expected benefits, some concerns exist that they may overload drivers' capacity and decrease performance. According to the multiple resource theory (Wickens, Hum Factors 50:449–455, <https://doi.org/10.1518/001872008X288394>, 2008), overload may occur at different stages of processing, that is, perceptual–central and/or response-related stages. Therefore, different measures may be needed to detect such specific demands. We explored the sensitivity of different mental workload measurements during the performance of an auditory task alone (single task) and in combination with a tracking task that was presented without (dual task) or, with a visual display (triple task). The demands associated with the *number of concurrent tasks* (single, dual and triple tasks), tracking *speed* (low, high, adjustable) and their interaction were analyzed. To account for different processing requirements, mental workload was assessed using subjective, behavioral (performance on the auditory task) and psychophysiological measurements (event-related potentials). 17 young adults participated in the study. The results showed that most measurements discriminated between the performances of one or more tasks, as well as between low and high speeds. However, only the subjective ratings and tracking task performance further discriminated between the dual- and triple-task conditions. Finally, ERPs (N1 and P3) were the only measure detecting increases in cognitive demands associated with higher requirements on processing speed combined with the addition of the display. Our results suggest that ERPs may provide complementary information to other traditional mental workload measures. Its applications in the evaluation and design of future systems should be investigated.

**Keywords** Mental workload · Event-related potentials · Processing resources · Time pressure · Cognitive demands and intelligent transport systems

## 1 Introduction

New intelligent transport systems (ITS) are entering the automotive market at a rapid pace. Some of these systems, such as the in-vehicle information systems (IVIS) or the advanced driver assistance systems (ADAS), directly interface with the drivers. IVIS collect and provide relevant information from the road, the vehicle or the environment to the driver by means of displays of different modalities (e.g., visual, auditory or haptic). Some examples are route guidance or hazard warning system. ADAS, on the other hand, assist

drivers in the control of certain vehicle dynamics (e.g., longitudinal and lateral control), or in the avoidance of potential collisions. Some examples are collision avoidance system, active cruise control or lane keeping assist. While IVIS and ADAS are intended to increase driving safety and comfort, concerns exist about their impact on driver demands and behavior.

In relation to this, the literature on ITS effects on driving has shown that such systems may negatively affect drivers' mental workload (MWL) by either, placing too high or too low demands on drivers' processing capacity. For example, different studies have reported some cognitive and behavioral effects when drivers are placed in a passive supervisory role during highly automated driving, such as underload, complacency or vigilance decrements (Young and Stanton 2002; Banks et al. 2018; Körber et al. 2015; Saxby et al. 2008). As a consequence, drivers have been shown to react

✉ I. Solís-Marcos  
ignacio.solis@vti.se

<sup>1</sup> Unit of Human Factors in the Transport System, Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden

more slowly when required to take over control (Eriksson and Stanton 2017; Gold et al. 2013; Larsson et al. 2014; Merat and Jamson 2009; Young and Stanton 2007) or to have a poorer awareness of the driving situation, among other effects (for a review, see de Winter et al. 2014). In contrast, IVIS have been shown to demand extra resources that may exceed drivers' limited processing capacity (Kahneman 1973), leading to an increased workload and/or distraction and performance decrements. For example, Reyes and Lee (2004) observed that a system presenting auditory information of the location of restaurants can increase drivers' brake reaction times. Similarly, the visual presentation of decision-making elements may increase drivers' MWL and decrease driving performance mostly reflected in a greater number of lane deviations (Blanco et al. 2006). In addition, driver's visual attention towards the road has been reported to be 'narrowed' when using either visual or non-visual IVIS due to the greater demands (Recarte and Nunes 2000; Strayer and Johnston 2001). In support of this, different studies have found evidence for decrements in the peripheral detection task (PDT) to assess the demands of IVIS of different modalities (Angell et al. 2006; Harms and Patten 2003; Merat and Jamson 2008).

This evidence underlines the need for detecting in time excessive increments or decrements in the driver MWL while interacting with one or various support systems that could negatively affect drivers' attention. For several reasons, this can be a difficult task. First, changes in the driver MWL may not show at first from the performance, and consequently, may go unnoticed until it is too late. For example, the underload effects of prolonged automated driving may only become evident when a take-over request is issued, and the driver's reaction is slower than anticipated. Second, drivers tend to compensate for high and also low demands. For example, additional demands due to cell-phone usage may be compensated for by decreasing driving speed (Patten et al. 2004). Likewise, boredom or fatigue induced by automated driving on a straight highway may be counteracted by engaging in non-driving-related tasks (e.g., reading, checking the phone), as suggested elsewhere (Atchley and Chan 2010; Neubauer and Matthews 2012). Such compensatory strategies need to be accounted for when exploring the actual effects of a specific system.

Furthermore, MWL is generally considered a multidimensional construct (Wickens 1984; Young et al. 2015), meaning that different factors could contribute to increasing or decreasing it, from 'internal' factors such as age or experience (e.g., Paxion et al. 2014; Cantin et al. 2009) to 'external' factors linked to the amount, type and complexity of the ongoing tasks (e.g., Baldwin et al. 2004; Patten et al. 2004). Concerning this, Wickens' multiple resource theory (Wickens 2008) posits that, rather than an undifferentiated pool of resources (Kahneman 1973), human processing

capacity consists of separate pools of resources along different dimensions: stage of processing (perceptual–central and response related), modality (visual or auditory) or code (verbal or spatial). The extent to which various tasks tap into similar resources will determine their level of interference. Thus, two tasks demanding perceptual–central resources will interfere more if one of them only requires response-related resources. Furthermore, different MWL measures do not always produce similar results (de Waard and Lewis-Evans 2014; Matthews et al. 2015; Yeh and Wickens 1988). This could indicate that different measures may be sensitive to different levels of task demands (de Waard 1996), but also to demands placed on different processing resources (Wickens 2008). For example, primary and secondary task performance indicators have been shown to detect high levels of MWL better than moderate or low levels (Wilson and Eggemeier 2006). In addition, performance decreases are usually the outcome of the complex interaction of different factors and do not inform about the nature of the demands and the resource pools that are overloaded. In this study, MWL will be assumed to be multidimensional and the multiple resource model (Wickens 2008) will be used as a framework to analyze the processing requirements placed by different factors on the different stages of processing.

Some psychophysiological techniques, such as the event-related potentials (ERPs), may complement the information provided by subjective and performance MWL indicators, as emphasized by Wickens (1990). ERPs consist of time-locked brain responses to specific events commonly presented in the context of a discrimination task in which an infrequent target signal needs to be detected in a series of noise signals (e.g., the so-called oddball task). An ERP is comprised of different temporally separated components characterized by three parameters: (a) the polarity of the component, positive (P) or negative (N), (b) the latency (usually expressed in milliseconds), which indicates 'when' the specific mental process represented by the component occurs (Kutas et al. 1977), and (c) the amplitude (in microvolts or  $\mu\text{V}$ ), which would reflect the 'intensity' of the processing, that is, the amount of neural resources allocated (Polich 2007). Based on this, early components like N1 (85–150 ms after stimulus onset) have been linked to processes involved in the 'perceptual' analysis of the stimuli (Kramer et al. 1983; Kok 1997), whereas later components like P3 have been associated with 'central' processes such as the semantic processing and categorization of the stimuli (Polich 2007). Specifically, N1 and P3 latencies and amplitudes elicited by a secondary task have been shown to be modulated by variations in the demands of primary complex tasks. Typically, longer latencies and lower amplitudes in different components have been reported when the primary task demands increase (Kramer et al. 1983; Allison and Polich 2008; Miller et al. 2011; Ullsperger et al. 2001). For this reason, ERPs may better

account for the specific processing requirements of complex tasks compared to other measures. Depending on which component (e.g., N1 and/or P3), and the parameter/s affected (i.e., latency and/or amplitude), it is possible to obtain a more comprehensive idea of where interferences occur and what resources are most demanded by the primary task.

Throughout the last four decades, several studies have explored the applicability of ERPs in operational tasks in laboratory settings (Allison and Polich 2008; Miller et al. 2011; Ullsperger et al. 2001), and to a lesser extent in ecological settings (e.g., Strayer et al. 2015). In transport research, ERPs have been used to evaluate the effects of different IVIS and other concurrent tasks while performing the primary task of driving. For example, Strayer et al. (2015) analyzed the cognitive effects of seven different in-vehicle tasks in three different settings: a laboratory setting, a driving simulator and an instrumented vehicle. It was observed that those in-vehicle tasks that required verbal and visual interaction (e.g., a speech-to-text e-mail system) significantly decreased P3 amplitude and the performance on a secondary detection task. In another set of studies, Baldwin and Coyne (2005) used a similar setting in which a detection task (i.e., an auditory or a visual oddball task) was used to detect changes in the primary task demands. The primary task consisted of a flight task to be performed while executing a single command (low-demand condition) or multiple commands (high-demand condition) presented via an information system in a visual or auditory modality. The P3 amplitude discriminated between the execution of the detection task alone or in combination with the flight task. However, it failed to detect gradual increases in MWL (i.e., from single to multiple commands conditions), which were better detected by the behavioral measurements. Also, contrary to the predictions of the multiple resource theory, no effects of intramodal combinations of the information system and oddball task were found, which is consistent with other studies (Baldwin and Coyne 2005; Merat and Jamson 2008).

The collected evidence to date shows that ERPs, and particularly P3 amplitude, discriminate when an operator is performing one or various tasks; however, its sensitivity to gradual increases in the primary task demand is not so clear. In many cases, the level of demand has been increased by manipulating the amount of information to be processed (e.g., Allison and Polich 2008; Baldwin and Coyne 2005). However, it is possible that, by itself, this is not enough to increase MWL. As posited by the time-based resource-sharing model (Barrouillet et al. 2004), the major factor for cognitive load is the available time to process the current cognitive demands. This has been later supported by Borragán et al. (2017), who showed that the available processing time influences cognitive load more than the complexity of the tasks itself. From this evidence it is inferred that, as long as an individual has or can make

available enough time to process all the concurrent tasks (e.g., texting while driving on a straight highway with low traffic), no great increases in cognitive demands will be experienced. Conversely, as the time pressure to process the tasks increases, so does the cognitive demand. Some well-established workload measures, like the NASA-TLX (Hart and Staveland 1988), assess temporal demand as one important contributing factor to the overall workload perceived by an operator. However, its specific effects on the operators' mental demands, and particularly on perceptual and cognitive stages of processing, have received less attention in the literature and, therefore, will be covered here.

The present study aimed at exploring the role of the amount of information and time pressure in the performance of a continuous tracking task while using a visual IVIS prototype. As in previous studies (Baldwin and Coyne 2005; Strayer et al. 2015) different MWL measurements were used, including ERPs. Such measurements were expected to inform about the mental demands placed by the different factors on different stages of processing described in the multiple resource theory (Wickens 2008). The amount of information factor was manipulated by comparing the performance of the auditory task alone, or simultaneously with a tracking task without and with the additional visual display. Time pressure was manipulated by changing the speed of the tracking task. Thus, potential interactions between the tracking speed and the presence or absence of display were investigated as well. A secondary objective of this study was to observe whether drivers strategically adjust the tracking speed to better cope with the ongoing demands, which should be reflected in MWL measures as well.

## 2 Materials and methods

### 2.1 Participants

We recruited 20 young adults to this study. However, three participants were removed for technical problems or too many artifacts in the brain signals. The mean age of the final sample ( $N = 17$ ) was 23.2, with a standard deviation of 4.06 years. They were required to refrain from consuming caffeine or tea for 4 h and alcohol for 24 h before the experiment day. All participants were right-handed and reported normal or corrected-to-normal visual acuity and hearing, as well as no history of neurological or psychiatric disease.

This study was carried out in compliance with the ethical guidelines of the Declaration of Helsinki and was approved by the Ethics Review Board from the University of Loyola Andalucía (Seville, Spain). All participants signed informed consent before their inclusion in the experiment.

## 2.2 Tracking task

A tracking computer task specifically designed for this study was used as the primary task. A one-lane road was presented on a computer screen with a small red car in the center of the lane (see Fig. 1). The scenario was presented from a bird's eye view and simulated a rural road with frequent mild curves. The participants' task was to keep the red car within the lane as much as possible using the joystick of an Xbox 360 controller. To investigate the effect of speed on MWL, the participants performed the task at three different speeds: low, high and adjustable. In the low- and high-speed conditions, the speed was fixed and could not be modified. The adjustable speed condition was included to investigate compensatory strategies for increased demands involving the tactical control of speed and their effects on the different MWL measurements. In such conditions, the speed was fixed, but it could be increased or decreased any time by means of two different buttons on the Xbox 360 controller. The low and high speeds were purposely chosen to create a noticeable difference between them. Given that the bird's view considerably reduced the perception of speed, the selected speed levels for the low- and high-speed conditions were unrealistically high. Therefore, for ease of interpretation, speeds were transformed to percentages using the speed level in the high-speed condition as the 100% reference, and no speed as the 0% reference. The chosen speed for the low-speed conditions was 66.6, as it represented the 66.6% of the high-speed condition. The standard deviation of the lateral position (SDLP), measured in meters, and the percentage of time off the track (% time off track) were used as performance indicators of the tracking task. Also, average speeds and speed variability were recorded when the speed was adjustable.

A prototype of an IVIS was designed that was relevant to the performance of the tracking task. It consisted of a visual continuous display (henceforth called 'preview display') that presented the simulated road from a higher up perspective.

On such display, participants could see their position (red dot on Fig. 1), and preview the upcoming road (equivalent to 1500–2000 meters in a real road). Thus, drivers could use such information to cognitively and/or behaviorally anticipate to the coming curves. However, given that the visual information presented by the display and the tracking task was very similar, it could happen that the preview display did not add any extra demand that could be detected by the MWL measures. To better simulate an IVIS that added extra demands, the 'preview display' was presented in a horizontal position. This way, the participants had to perform mental rotations to match the visual information from both sources, which expectedly would increase driver demands. Drivers were instructed to use the preview display as much as possible when it was available.

## 2.3 MWL measurements: subjective MWL, auditory oddball task and ERP

### 2.3.1 Subjective MWL

As described earlier, the focus of this work was to detect and analyze the effects of the amount of information and time pressure on the participants' mental demands, namely on perceptual and cognitive resources. For this purpose, subjective mental workload was assessed after each condition by asking the participants to rate the question "how mentally demanding was the task?" (obtained from the "Mental demand" sub-scale from the NASA-TLX; Hart and Staveland 1988) on a scale from 0 (low) to 20 (high). Scores obtained from the participants were transformed into a 0–100 scale for ease of interpretation.

### 2.3.2 Auditory oddball task

An auditory oddball task was used as secondary task to elicit the ERPs. Similar tasks have been previously used in other studies analyzing the effects of IVIS on operators' MWL

**Fig. 1** Computer tracking task with the preview display on top of the screen. In the preview display, the ego-car is represented by a red dot. (Colour figure online)



(e.g., Baldwin and Coyne 2005; Strayer et al. 2015). In each condition, a sequence of high-pitch tones (target, 1200 Hz) and low-pitch tones (standard, 800 Hz) sounds was randomly presented binaurally via headphones (70 dB) at fixed intervals of 1 s. Each tone was presented for 50 ms and the probability ratio was 80% for the standard tones and 20% for the target tones. The duration of each block (1 per condition) was 5–6 min, approximately. The sequence of the tones was randomized across participants and conditions. Participants were instructed to respond as quickly and accurately as possible to the target tones by pressing the button attached on the back side of the controller.

### 2.3.3 ERP: recording and procedure

Electrical brain activity was continuously recorded from 58 scalp electrodes distributed according to the 10–20 international system and compared to an averaged reference (Fp1, Fpz, Fp2, F3A, F4A, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FC5, FC3, FC1, FCz, FC2, FC4, FC6, T3, C5, C3, C1, Cz, C2, C4, C6, T4, T3L, CP5, CP3, CP1, CPz, CP2, CP4, CP6, T4L, T5, P5, P3, P1, Pz, P2, P4, P6, T6, PO5, PO3, PO1, POz, PO2, PO4, PO6, O1, Oz, and O2). Horizontal and vertical ocular electrodes (HEOG and VEOG, respectively) were placed on the outer left and right canthi and above and below the left eye. Impedance levels were kept below 5 kOhms. The electrode signals were amplified with BrainAmp amplifiers and digitally stored using Brain Vision Recorder software (Brain Products GmbH, Germany).

EEG data analysis was performed in Matlab R2014b using EEGLab 13.4.4b, an open-source toolbox developed by Delorme and Makeig (2004). Data were sampled at 500 Hz and filtered offline with a 0.1-Hz high-pass and a 50-Hz low-pass filter. ERP averages were calculated from 200 ms prior to and 800 ms after the stimulus presentation. All ERP averages were corrected to the baseline over the pre-stimulus interval (–200 to 0 ms). Artifacts related to muscle, blinks and ocular movements were minimized using the ‘runica’ ICA algorithm (Lee et al. 1999). Besides, a semi-automated identification of eye-related artifacts was conducted using the ADJUST plug-in for EEGLab (Mognon et al. 2010). Moreover, those trials where the amplitude exceeded  $\pm 75 \mu\text{V}$  on HEOG electrode were discarded. All recordings with less than 40 epochs valid were excluded from the analysis to guarantee an adequate signal-to-noise ratio.

N1 and P3 components were subtracted from each participant and condition following Duncan et al.’s guidelines (2009), that is, for each component, the electrode showing the maximum amplitude and its exact latency were identified. The different components were identified in the following time windows after stimulus onset: N1 = 80–150 ms and P3 = 270–400 ms.

## 2.4 Experimental design

A  $2 \times 3$  within-subject design with two independent variables was used. One independent factor was *preview display*, which consisted of two levels: off (or dual task), and on (or triple task). The other independent factor was *speed*, which consisted of three levels: ‘low’, ‘high’ and ‘adjustable’. In addition to these six conditions, the participants performed a ‘view’ condition (or single task condition) in which they performed the auditory oddball task while fixated on the paused tracking task on the screen. The purpose of adding this condition was to explore which MWL measurements were the most sensitive to the additional of the tracking task and the visual display to the auditory oddball. All conditions were randomized and each participant was presented with a different order.

## 2.5 Procedure

The present study took place in a physiology laboratory equipped with an electromagnetically isolated room. Upon arrival, the participant was briefed with general information about the different tasks to be performed. Then, an informed consent sheet was signed in case they agreed with all the conditions.

Next, the participant was guided to a shielded chamber and seated approximately 65 cm from a computer screen. Then, he/she was given practice on the oddball and the tracking task, first separately and then simultaneously. The total training time lasted approximately 15 min. After the completion of the training, the 58-electrode EEG equipment was set on the participant.

Afterwards, the participant was guided back to the isolated room to perform all the different conditions in a semi-counterbalanced order. The participant was instructed to keep the ego-vehicle within the lane while responding to the target tone as quickly and accurately as possible. Also, he/she was instructed to use the preview display as much as possible. Each condition started with 15 s of only tracking task to let the participant familiarize with the low or high tracking speeds, and to find a comfortable speed in the adjustable speed conditions. No data were logged during this period. Physiological recordings were logged throughout the whole drive. Once all the conditions were completed, all the electrodes were removed and the participant was thanked for his/her participation.

## 2.6 Statistical analyses

The different dependent variables included in our analyses are summarized in Table 1. As shown, the types of MWL

**Table 1** Dependent variables (MWL measures) included in our analyses

Processing stages (multiple resource theory)	Type of MWL measurements	Dependent variables
Perceptual	ERPs	N1 latency N1 amplitude
Central		P3 latency P3 amplitude
Response	Auditory task	Accuracy Reaction time
	Tracking task	% time off the track SDLP
-	Subjective MWL	Mental demand

measurements are categorized according to their expected sensitivity to the different stages of processing described in the multiple resource theory (Wickens 2008).

Means and standard errors (SE) for each dependent variable in each condition were calculated. Parametricity of most of the variables was confirmed using Shapiro–Wilk’s test.

Two separate analyses of variance were conducted to explore the effects of *number of concurrent task*, *display* and *speed*:

1. To analyze the effect of the *number of concurrent tasks* on the different MWL measurements, one-way independent ANOVAs with subjects as random factor were conducted comparing the three conditions: (a) auditory task alone (single task), (b) auditory task and tracking task simultaneously (dual task) and (c) auditory task, tracking task and display simultaneously (triple task). Post hoc comparisons using Bonferroni’s correction were performed to compare the three levels. The specific post hoc comparisons between the dual and triple tasks were used to report the effects of *display*.
2. *Speed* effects and interactions with the *preview display* variable were further explored using  $2 \times 3$  repeated measures ANOVAs. Greenhouse–Geisser’s correction was applied when Mauchly’s test detected violations of sphericity. Bonferroni’s corrections were used to protect against type-I error.

Alpha level was set at 0.05. The partial eta-square was used as a measure of relative effect size.

### 3 Results

In this section, the results obtained from the different analyses will be presented to respond to our different objectives:

- (a) Effects of *number of concurrent tasks* on subjective MWL, auditory task performance and ERPs. Also, the *display* effects on the primary task performance were included here.
- (b) Main effects of *speed* on all MWL measures and tracking task performance.
- (c) Interaction effects of *speed* and *display* on all MWL measures and tracking task performance.

A summary table is provided with the main and interaction effects observed in the different variables analyzed (see Table 2).

#### 3.1 Effects of number of concurrent tasks

Our analyses revealed significant effects of *number of concurrent tasks* (single, dual and triple task) on different dependent variables including subjective MWL, accuracy, reaction times, N1 amplitude, P3 amplitude and P3 latency (see Table 2). Post hoc analysis of these variables was consistent in showing significant differences between the single task and the other two conditions (all  $P$  values  $< 0.01$ ). All scores and ERP values indicate that MWL was lowest in the single-task condition (see Fig. 2; Table 4). Only subjective MWL showed further differences between the dual- and triple-task conditions. Higher demands were perceived as more tasks were required to be performed (Fig. 2). An illustration of the grand average waveforms of N1 and P3 in the three conditions is presented in Fig. 3a, b, respectively.

Additionally, the preview display effect on the tracking task performance was explored by comparing the dual- and triple-task conditions. The results showed that the participants spent a greater percentage of time off the track when the display was present (see Table 3).

#### 3.2 Effects of speed

Specific analyses were carried out to observe whether, when allowed to adjust speed, the participants preferred different speeds than the predefined levels for the low- (66.6%) and high-speed conditions (100%). Table 3 shows the mean tracking speeds in percentages (0% = no speed and 100% = speed in the high-speed conditions). As shown, in the adjustable speed conditions, the drivers selected speeds that were above the 90% of the highest speed. Pairwise comparisons were conducted after a significant repeated measures ANOVA was found ( $F(2, 32) = 43.97, P < .01, \eta_p^2 = 0.73$ ). No differences between the adjustable and high-speed conditions were found, but significant differences between these and the low-speed condition were observed (both  $P$  values  $< 0.01$ ).

As shown in Table 3, the tracking speed affected the behavioral, subjective and physiological measurements. All

**Table 2** A summary table with the significant found effects

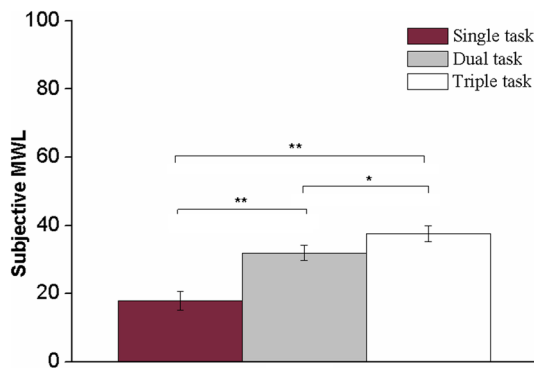
Processing stages (multiple resource theory)	Type of MWL measurements	Dependent variables	Number of tasks $F(2,32)$	Speed $F(2,32)$	Speed × preview display $F(2,32)$	Direction of the significant effects
Perceptual	ERPs	N1 latency			$F = 4.4$ $\eta_p^2 = .22$	Dual < triple (in adjustable and high speed) <sup>b</sup>
		N1 amplitude	<b><math>F = 6.8</math></b> $\eta_p^2 = .29$	$F = 3.3$ $\eta_p^2 = .17$		Dual and triple < single <sup>a</sup> high and adjustable < low <sup>b</sup>
Central		P3 latency	<b><math>F = 20.56</math></b> $\eta_p^2 = .56$			Dual and triple < single <sup>a</sup>
		P3 amplitude	<b><math>F = 15</math></b> $\eta_p^2 = .49$		$F = 3.77$ $\eta_p^2 = .19$	Dual and triple < single <sup>a</sup> triple < dual (in low speed) <sup>b</sup>
Response	Auditory task	Accuracy	<b><math>F = 20.56</math></b> $\eta_p^2 = .56$	<b><math>F = 7.98</math></b> $\eta_p^2 = .32$		Dual and triple < single <sup>a</sup> high < low and adjustable <sup>b</sup>
		Reaction time	<b><math>F = 33.87</math></b> $\eta_p^2 = .68$	<b><math>F = 6.33</math></b> $\eta_p^2 = .27$		Dual and triple < single <sup>a</sup> low < high and adjustable <sup>b</sup>
	Tracking task	% time off the track	<b><math>F = 16.05</math></b> $\eta_p^2 = .49$	<b><math>F = 11.03</math></b> $\eta_p^2 = .39$		Dual < triple <sup>a</sup> low < high and adjustable <sup>b</sup>
		SDLP			<b><math>F = 9.63</math></b> $\eta_p^2 = .36$	Low < high and adjustable <sup>b</sup>
	–	Subjective MWL	Mental demand	<b><math>F = 14.07</math></b> $\eta_p^2 = .47$	<b><math>F = 9.28</math></b> $\eta_p^2 = .55$	

Single single task (only auditory oddball), Dual dual task (auditory oddball+tracking task), Triple triple task (auditory oddball+tracking task + display)

F values and effect sizes are provided. The significance level is represented by fonts, \* $p < .05$  (italics); \*\* $p < .01$  (bold italics)

<sup>a</sup>Post hoc analysis corresponding to the independent one-way ANOVA

<sup>b</sup>Post hoc analysis corresponding to the 2 × 3 repeated measures ANOVA



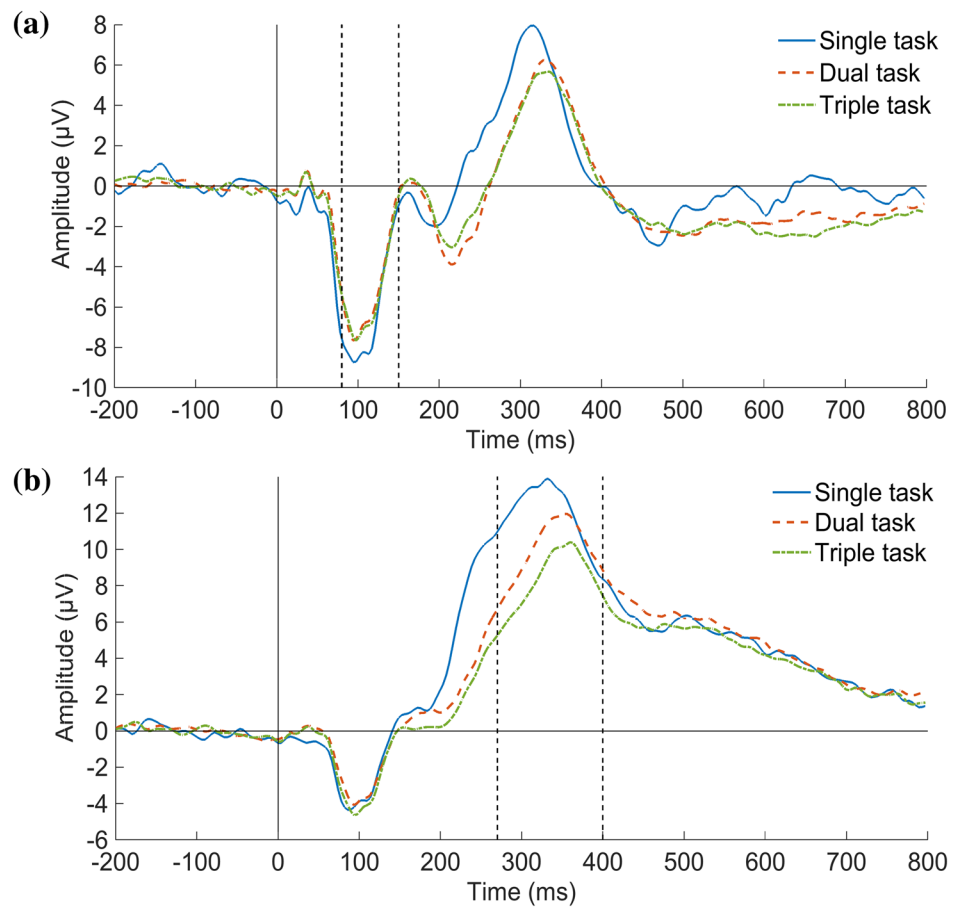
**Fig. 2** Subjective MWL global scores in the single-, dual- and triple-task conditions

post hoc analyses were consistent in showing significant differences between the low-speed conditions and the high- and adjustable speed conditions. When the tracking speed was low, the participants performed better in the tracking task (as shown by lower lateral variabilities and more time in the track) and auditory task (higher accuracy and shorter response times). Also, they reported lower subjective MWL and showed larger N1 amplitudes (Table 4; Fig. 4).

### 3.3 Interaction effects between speed and preview display

Interaction effects were observed only in the ERP measures, but not in the other MWL measures (see Table 2). N1 latency showed increments in the high and adjustable conditions when the visual display was on ( $P < .05$ ), but similar

**Fig. 3** Global waveforms of N1 (a) and P3 (b) components in the single, dual and triple tasks. The time windows for each component are represented with dotted lines. *Fz* and *Pz* electrodes were used to represent N1 and P3 waveforms, respectively



**Table 3** Performance indicators from the tracking task

Conditions		% time off the track		SDLP		% speed	
Preview display	Speed	Mean	SE	Mean	SE	Mean	SE
ON	Low	19.33	2.53	1.18	0.07	66.6	–
	Adjustable	21.59	2.59	1.31	0.08	93.43	5.1
	High	23.24	3.01	1.36	0.09	100	–
OFF	Low	14.34	1.91	1.08	0.05	66.6	–
	Adjustable	20.39	3.33	1.45	0.15	90.22	6
	High	22.83	2.47	1.31	0.07	100	–

All means and SEs are presented except for % speed in the low- and high-speed conditions in which the speed was fixed

values were found in the low-speed condition. In addition, P3 amplitude differences were observed only between the low-speed conditions, with higher amplitudes when the display was off ( $P < .05$ ) (see Fig. 5).

## 4 Discussion

For a better understanding of how and where in the information processing chain different aspects of task demand affect mental workload, we compared different mental workload

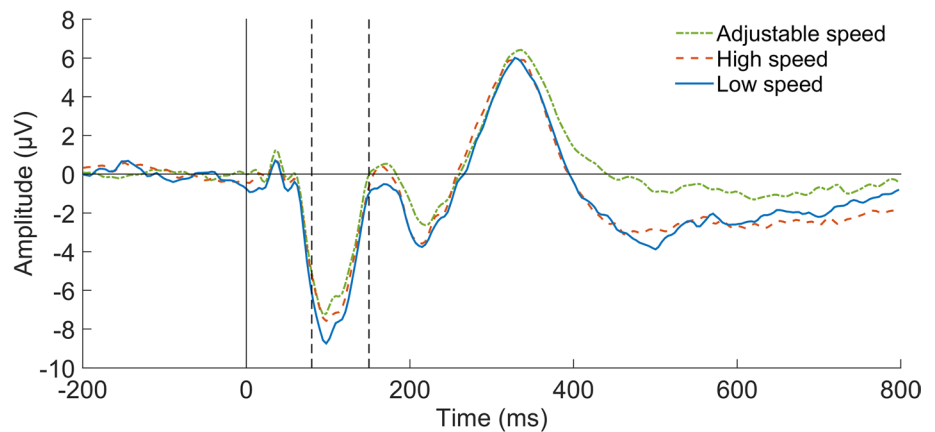
measures, tapping into different stages of information processing. Performance decrements can be seen as indicators showing effects at the end of the chain. If they occur, then no compensatory strategies further up the chain prevented the decrements; however, it is not clear, where along the chain the problems occurred. As shown previously, performance and subjective workload measures are not directly related (Vidulich and Wickens 1986; Yeh and Wickens 1988), that is, performance decrements can occur without a perceived increase in workload, as well as increased workload can be perceived without showing effects in performance. Given



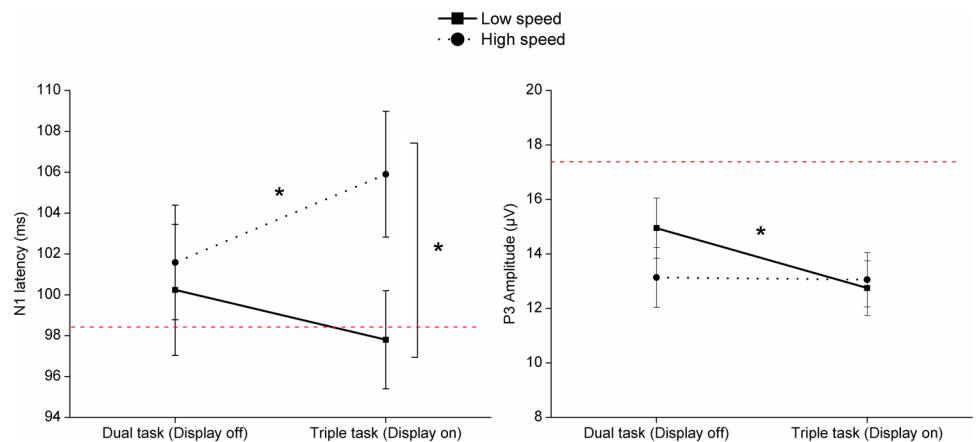
**Table 4** Subjective and performance measurements from the auditory task

Conditions		Subjective MWL		Accuracy on target		Reaction time	
		Mean	SE	Mean	SE	Mean	SE
	View	17.94	2.79	96.11	1.05	305.63	8.54
Preview display	Speed	Mean	SE	Mean	SE	Mean	SE
On	Low	30.88	3.07	86.47	2.09	347.65	7.33
	Ajustable	40.29	3.7	85.25	2.98	362.95	8.03
	High	43.82	4.4	80.58	3.11	358.99	9.68
Off	Low	27.35	3.97	88.7	1.8	347.85	8.14
	Adjustable	34.41	3.48	85.17	2.43	366.2	9.9
	High	35.29	3.65	78.8	4.27	364.7	8.7

**Fig. 4** Main effect of speed on N1 amplitude. N1 amplitude in the low-speed condition was greater than that in the other two conditions



**Fig. 5** Interaction effects observed in N1 latency (on the left) and P3 amplitude (on the right). Values from the single-task condition are represented by the red dotted line. The adjustable speed condition was not included as it was similar and had the effect. (Colour figure online)



this dissociation, it can also be speculated that workload can increase, without immediate effects on either performance or subjective perception of load. Especially here lies the potential of ERPs, as they may detect changes in processing ability early on.

Overall, our results show that while some MWL measures are sensitive to the demands placed by one specific factor (e.g., number of tasks), others may detect various

sources of demands or their interactions. This supports the idea suggested by other authors (e.g., de Waard and Lewis-Evans 2014; Baldwin and Coyne 2005) that different MWL measurements are necessary to account for different sources of demand during dynamic multitasking, such as driving while using an IVIS. In addition, our findings showed that ERPs can identify where in the information processing chain different demands take effect. Adding a continuous visual

tracking task to the auditory oddball task showed in all available measures, corroborating the results from other studies, that additional processing requirements put more strain on perception and central processing, as shown by the ERPs, eventually leading to performance decrements as compared to single-task performance. Here, the additional demand was also experienced and reported subjectively. Further visual and cognitive load from adding the preview display, which did not require an active response, but a cognitive integration of two visual information sources, affected task performance on the tracking task and was perceived as demanding, but did not disturb performance on the auditory task. This might be explained by the auditory task and visual display relying on different sensory modalities and that they are not overloading common cognitive resources.

Increased demands in information processing speed affected attention allocation (N1) and multiple task performance, resulting in an increased subjective workload. Most employed workload measures concur in indicating higher loads for higher time pressure, operationalised by a higher speed in the tracking task. This generalized effect on all measurements shows that external demands on processing speed are not specific to one type of resource. Instead, speed modulates the amount of resources needed from the different resource pools used for different tasks. A higher speed in the tracking task increases demands, as the curve frequency and the required response rate increases. This has direct detrimental effects on the performance in the tracking task. Performance also decreased in the auditory task, indicating that an increase in speed affects attentional processes involved in the selection of the target tones. This is consistent with the observed decrements in N1 amplitude, a parameter that has commonly been linked to early attentional selection processes and allocation of perceptual resources (Kramer et al. 1983; Kok 1997). Congruently, the participants reported a perceived higher workload.

When combining the time-based resource-sharing model (Barrouillet et al. 2004) and the multiple resource theory (Wickens 2008), the highest MWL is expected for the high-speed condition with the visual display on. The performance measures and subjective ratings are in line with these expectations. The N1 latency and P3 amplitude, known to be valid components to assess demands on a perceptual–central level (Kramer and Spinks 1991; Polich 2007), and to provide more detailed information about the actual processes occurring, show that the different types of demands interact in their effect on workload. The mobilization of perceptual resources (N1 latency) is only affected in the triple task condition, when the requirements on processing speed are high. At a low processing speed, neither the addition of the tracking task nor further adding the display leads to a significant decrement, while at a high processing speed, one additional visual task can be handled, but not two. The availability of

central resources to process the auditory task (P3 amplitude), on the other hand, is already affected by adding the tracking task. At a low processing speed, adding the display as well leads to a measurable decrease in resource availability. For high demands on processing speed, the P3 amplitude immediately decreases already in the dual-task condition to the triple-task level in the low-speed condition. This can either imply that the P3 amplitude reached a “floor” level, or that triple-task performance is extraordinarily resource demanding, regardless of the available processing speed. The latter would imply that there is a risk to overload operators with “too much information” at the same time, regardless of whether they actually have more time available to process the information.

Overall, these findings seem to indicate that ERP is a suitable technique to detect increases in cognitive demands associated with interactions between the processing speed and amount of information to be processed. However, such interaction affected N1 and P3 components in different ways. The effects in N1 latency were only evident when demands were probably highest, that is, when combining the visual display with high processing speeds. By contrast, the effects of the display on P3 amplitude were only detectable when the demands were low to moderate. This could show that, only at low processing demands without additional display, the participants were able to allocate more resources to the auditory task. These findings might reflect that cognitive resources are more sensitive to lower demands of the interaction and more insensitive to increased demands. Perceptual resources, on the other hand, would show the opposite pattern. Based on this interpretation, it would be recommendable to analyze more than one ERP component to better understand the interaction effects of processing speed and number of ongoing tasks.

While the use of ERPs to analyze MWL in complex settings needs to be further investigated, we could speculate about its potential applications in the evaluation and design of more ergonomic ITS. For example, ERPs may guide the integration of various IVIS in one vehicle while ensuring the lowest interference level possible between them and with the driving task. In particular, ERPs could be used to better highlight conflicts between systems that require similar perceptual and/or cognitive resources. Additionally, in the light of our findings, ERPs could also be used to design systems that are ‘aware’ of drivers’ time pressure to process the ongoing tasks. For example, future adaptive automation systems may dynamically adjust the moment, the amount and/or the modality of the information presented when the time pressure is expected to be higher. Furthermore, ERPs may also be a sensitive method to detect the attentional reductions associated with the potential underload effects of automated driving (Young and Stanton 2002; Saxby et al. 2008). Given that in a highly automated driving, driver intervention

is expected to be required only in some occasions (e.g., when the system fails), detecting his/her previous readiness to react in a timely and safe manner would be very recommendable. In this regard, ERPs could then help to determine whether the driver attentional state is optimal to handle such potential situations or whether specific warning systems are necessary to raise their alertness level.

#### 4.1 Limitations of the study

In the high- and low-speed conditions, the speed and, therefore, processing demand was set by the experimenter, but in the adjustable condition the participants were given the possibility to influence processing demands via speed adjustment. We hypothesized that the participants would tactically adjust the speed to a comfortable level of task demand, which was expected to be somewhere between the preset high and low speed. However, the participants voluntarily chose the same speed as in the high-speed condition, which is why the workload measures do not differ between the two conditions either. Consequently, the expected beneficial effects of managing speed on MWL could not be explored. This behavior contradicts what has been observed in other studies, which show compensatory decrements in speed in demanding conditions (e.g., Patten et al. 2004; Törnros and Bolling 2006; Ahlstrom et al. 2016). A potential explanation could be that the participants chose higher speeds to fight boredom, accepting a performance decrease and increased mental workload, as this was not associated with any real costs. In real traffic, decreased driving performance can have harsh consequences, which could be simulated with an adjusted pay-off matrix in future laboratory tasks.

## 5 Conclusions

Our results indicate that, in multitasking conditions, such as driving while using information systems, it is highly recommendable to use different MWL measurements to account for different sources of demand. As observed here, different measurements converged in the detection of similar types of demands, such as increases in processing speed. However, some dissociations were also observed. For example, subjective MWL was sensitive to the addition of the preview display, whereas ERPs detected interaction effects between the amount of information and processing speeds. A more comprehensive evaluation and design of future information systems may be achieved using different techniques. ERP applications should be further investigated as they could provide complementary information to other more traditional MWL measurements, particularly by detecting demands placed on perceptual and cognitive resources.

**Acknowledgements** This research was funded by the project HFAuto—Human Factors of Automated Driving (PITN-GA-2013-605817). We are grateful to Alejandro Galvao Carmona and Manuel Vázquez Marufo for their valuable contribution to improve this paper.

#### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Ahlstrom C, Kircher K, Thorslund B, Adell E (2016) Bicyclists' visual strategies when conducting self-paced vs. system-paced smart-phone tasks in traffic. *Transp Res Part F Traffic Psychol Behav* 41:204–216. <https://doi.org/10.1016/j.trf.2015.01.010>
- Allison BZ, Polich J (2008) Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biol Psychol* 77:277–283. <https://doi.org/10.1016/j.biopsycho.2007.10.014>
- Angell L, Auflick J, Austria PA, Kochhar D (2006) Driver workload metrics Project: final report (DOT HS 810 635)
- Atchley P, Chan M (2010) Potential benefits and costs of concurrent task engagement to maintain vigilance: a driving simulator investigation. *Hum Factors J Hum Factors Ergon Soc* 53:3–12. <https://doi.org/10.1177/0018720810391215>
- Baldwin CL, Coyne JT (2005) Dissociable aspects of mental workload: examinations of the p300 erp component and performance assessments. *Psychologia* 48:102–119. <https://doi.org/10.2117/psysoc.2005.102>
- Baldwin CL, Freeman FG, Coyne JT (2004) Mental workload as a function of road type and visibility: comparison of neurophysiological, behavioral, and subjective indices. *Proc Hum Factors Ergon Soc Annu Meet* 48:2309–2313. <https://doi.org/10.1177/154193120404801922>
- Banks VA, Eriksson A, O'Donoghue J, Stanton NA (2018) Is partially automated driving a bad idea? Observations from an on-road study. *Appl Ergon* 68:138–145. <https://doi.org/10.1016/j.apergo.2017.11.010>
- Barrouillet P, Bernardin S, Camos V (2004) Time constraints and resource sharing in adults' working memory spans. *J Exp Psychol Gen* 133:83–100. <https://doi.org/10.1037/0096-3445.133.1.83>
- Blanco M, Biever WJ, Gallagher JP, Dingus TA (2006) The impact of secondary task cognitive processing demand on driving performance. *Accid Anal Prev* 38:895–906. <https://doi.org/10.1016/j.aap.2006.02.015>
- Borragán G, Slama H, Bartolomei M, Peigneux P (2017) Cognitive fatigue: a time-based resource-sharing account. *Cortex*. <https://doi.org/10.1016/j.cortex.2017.01.023>
- Cantin V, Lavallière M, Simoneau M, Teasdale N (2009) Mental workload when driving in a simulator: effects of age and driving complexity. *Accid Anal Prev* 41:763–771. <https://doi.org/10.1016/j.AAP.2009.03.019>
- de Waard D (1996) The measurement of drivers' mental workload. Doctoral dissertation, University of Groningen, Haren

- de Waard D, Lewis-Evans B (2014) Self-report scales alone cannot capture mental workload. *Cogn Technol Work* 16:303–305. <https://doi.org/10.1007/s10111-014-0277-z>
- de Winter JCF, Happee R, Martens MH, Stanton N (2014) Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. *Transp Res Part F Traffic Psychol Behav* 27:196–217. <https://doi.org/10.1016/j.trf.2014.06.016>
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Duncan CC, Barry RJ, Connolly JF et al (2009) Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin Neurophysiol* 120:1883–1908. <https://doi.org/10.1016/j.clinph.2009.07.045>
- Eriksson A, Stanton NA (2017) Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Hum Factors J Hum Factors Ergon Soc* 59:689–705. <https://doi.org/10.1177/0018720816685832>
- Gold C, Dambock D, Lorenz L, Bengler K (2013) “Take over!” How long does it take to get the driver back into the loop? *Proc Hum Factors Ergon Soc Annu Meet* 57:1938–1942. <https://doi.org/10.1177/1541931213571433>
- Harms L, Patten C (2003) Peripheral detection as a measure of driver distraction. A study of memory-based versus system-based navigation in a built-up area. *Transp Res Part F Traffic Psychol Behav* 6:23–36. [https://doi.org/10.1016/S1369-8478\(02\)00044-X](https://doi.org/10.1016/S1369-8478(02)00044-X)
- Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv Psychol* 52:139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Kahneman D (1973) *Attention and effort*. Prentice Hall, Englewood Cliffs
- Kok A (1997) Event-related-potential (ERP) reflections of mental resources: A review and synthesis. *Biol Psychol* 45:19–56. [https://doi.org/10.1016/S0301-0511\(96\)05221-0](https://doi.org/10.1016/S0301-0511(96)05221-0)
- Körber M, Cingel A, Zimmermann M, Bengler K (2015) Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manuf* 3:2403–2409. <https://doi.org/10.1016/j.promfg.2015.07.499>
- Kramer A, Spinks J (1991) Capacity views of human information processing. In: Jennings JR, Coles MGH (eds) *Handbook of cognitive psychophysiology: central and autonomic nervous system approaches*. Wiley, New York, pp 179–242
- Kramer AF, Wickens CD, Donchin E (1983) An analysis of the processing requirements of a complex perceptual-motor task. *Hum Factors* 25:597–621
- Kutas M, McCarthy G, Donchin E (1977) Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197:792–795. <https://doi.org/10.1126/science.887923>
- Larsson AFL, Kircher K, Andersson Hultgren J (2014) Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving. *Transp Res Part F Traffic Psychol Behav*. <https://doi.org/10.1016/j.trf.2014.05.008>
- Lee TW, Girolami M, Sejnowski TJ (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput* 11:417–441. <https://doi.org/10.1162/089976699300016719>
- Merat N, Jamson AH (2008) The effect of stimulus modality on signal detection: implications for assessing the safety of in-vehicle technology. *Hum Factors* 50:145–158. <https://doi.org/10.1518/001872008X250656>
- Merat N, Jamson AH (2009) How do drivers behave in a highly automated car? In: *Proceedings of the Fifth International Driving Symposium on Human Factor in Driver Assessment, Training and Vehicle Design*, Montana, USA, pp 514–521
- Matthews G, Reinerman-Jones LE, Barber DJ, Abich J (2015) The Psychometrics of Mental Workload. *Hum Factors J Hum Factors Ergon Soc* 57:125–143. <https://doi.org/10.1177/0018720814539505>
- Miller MW, Rietschel JC, McDonald CG, Hatfield BD (2011) A novel approach to the physiological measurement of mental workload. *Int J Psychophysiol* 80:75–78. <https://doi.org/10.1016/j.ijpsycho.2011.02.003>
- Mognon A, Jovicich J, Bruzzone L, Buiatti M (2010) ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. <https://doi.org/10.1111/j.1469-8986.2010.01061.x>
- Neubauer C, Matthews S, Saxby D (2012) The effects of cell phone use and automation on driver performance and subjective state in simulated driving. In: *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting: October 22–26, 2012*, Boston, pp 1987–1991
- Patten CJ, Kircher A, Östlund J, Nilsson L (2004) Using mobile telephones: cognitive workload and attention resource allocation. *Accid Anal Prev* 36:341–350. [https://doi.org/10.1016/S0001-4575\(03\)00014-9](https://doi.org/10.1016/S0001-4575(03)00014-9)
- Paxion J, Galy E, Berthelon C (2014) Mental workload and driving. *Front Psychol* 5:1–11. <https://doi.org/10.3389/fpsyg.2014.01344>
- Polich J (2007) Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118:2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Recarte MA, Nunes LM (2000) Effects of verbal and spatial-imagery tasks on eye fixations while driving. *J Exp Psychol Appl* 6:31–43. <https://doi.org/10.1037/1076-898X.6.1.31>
- Reyes ML, Lee JD (2004) The influence of IVIS distractions on tactical and control levels of driving performance. *Proc Hum Factors Ergon Soc Annu Meet* 48:2369–2373. <https://doi.org/10.1177/154193120404801935>
- Saxby DJ, Matthews G, Hitchcock EM, Warm JS, Funke GJ, Gantzer T (2008) Effect of active and passive fatigue on performance using a driving simulator. In: *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting: September 22–26, 2008*, Santa Monica, pp 1751–1755
- Strayer DL, Johnston WA (2001) Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone. *Psychol Sci* 12:462–466. <https://doi.org/10.1111/1467-9280.00386>
- Strayer DL, Turrill J, Cooper JM et al (2015) Assessing cognitive distraction in the automobile. *Hum Factors* 57:1300–1324. <https://doi.org/10.1177/0018720815575149>
- Törnros J, Bolling A (2006) Mobile phone use—effects of conversation on mental workload and driving speed in rural and urban environments. *Transp Res Part F Traffic Psychol Behav* 9:298–306. <https://doi.org/10.1016/j.trf.2006.01.008>
- Ullsperger P, Freude G, Erdmann U (2001) Auditory probe sensitivity to mental workload changes—an event-related potential study. *Int J Psychophysiol* 40:201–209. [https://doi.org/10.1016/S0167-8760\(00\)00188-4](https://doi.org/10.1016/S0167-8760(00)00188-4)
- Vidulich MA, Wickens CD (1986) Causes of dissociation between subjective workload measures and performance. Caveats for the use of subjective assessments. *Appl Ergon* 17:291–296. [https://doi.org/10.1016/0003-6870\(86\)90132-8](https://doi.org/10.1016/0003-6870(86)90132-8)
- Wickens CD (1984) Processing resources in attention. In: Parasuraman R and Davies R (eds) *Varieties of attention*, Academic Press, London, pp 63–102
- Wickens CD (1990) Applications of event-related potential research to problems in human factors. In: Rohrbaugh JW, Parasuraman R, Johnson R, Jr. (eds) *Event-related brain potentials: Basic issues and applications*. Oxford University Press, London, pp 301–309
- Wickens CD (2008) Multiple resources and mental workload. *Hum Factors J Hum Factors Ergon Soc* 50:449–455. <https://doi.org/10.1518/001872008X288394>

- Wilson GF, Eggemeier FT (2006) Mental workload measurement. In: Karwowski W (ed) International encyclopedia of ergonomics and human factors, 2nd edn. CRC Press, Boca Raton
- Yeh Y, Wickens CD (1988) Dissociation of performance and subjective measures of workload. *Hum Factors J Hum Factors Ergon Soc* 30:111–120. <https://doi.org/10.1177/001872088803000110>
- Young MS, Stanton NA (2002) Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Hum Factors J Hum Factors Ergon Soc* 44:365–375. <https://doi.org/10.1518/0018720024497709>
- Young MS, Stanton NA (2007) Back to the future: brake reaction times for manual and automated vehicles. *Ergonomics* 50:46–58. <https://doi.org/10.1080/00140130600980789>
- Young MS, Brookhuis K, Wickens CD, Hancock P (2015) State of science: mental workload in ergonomics. *Ergonomics* 139:1–17. <https://doi.org/10.1080/00140139.2014.956151>