Transportation Research Part F 14 (2011) 199–208

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Driver workload and eye blink duration

Simone Benedetto [a], Marco Pedrotti [a,*], Luca Minin [b], Thierry Baccino [c], Alessandra Re [a], Roberto Montanari [b]

[a] *LIDEA - InterDepartmental Laboratory of Applied Ergonomics, Università di Torino, via Verdi 10, 10124 Torino, Italy*
[b] *Department of Science and Engineering Methods, Università di Modena e Reggio Emilia, via Amendola 02, 42122 Reggio Emilia, Italy*
[c] *LUTIN (UMS-CNRS 2809), Cité des sciences et de l'industrie de la Villette, Avenue Corentin Cariou 30, 75930 Paris, France*

ARTICLE INFO

Article history:
Received 31 July 2009
Received in revised form 3 December 2010
Accepted 3 December 2010

Keywords:
Driver workload
Blink duration
Lane Change Test

ABSTRACT

The use of standardized methods in driver distraction research is essential for comparing results across studies. This work examined the effects of in-vehicle information systems (IVIS) usage on eye blinks in a simulated Lane Change Test (LCT), a simple driving task specifically designed by the International Organization for Standardization. Fifteen participants performed the LCT in a driving simulator in both single- and dual-task conditions, the latter manipulated by introducing an IVIS task in the car cockpit. Results suggest that blink duration (BD), with respect to blink rate (BR), is a more sensitive and reliable indicator of driver visual workload. Besides considering mean BD values, a detailed analysis revealed that the distribution of BD follows a Gaussian-like curve in normal driving conditions: three duration classes (short, medium, long) were extracted from such distribution, and changes happening to each class were analyzed within the dual-task conditions. Short and long blinks reflect, respectively, the effects of visual workload and time on task: more short blinks occur with an IVIS interaction during driving, while more long blinks arise as time spent driving increases. These results may have practical implications for system design in automotive.

1. Introduction

The use of a wide variety of in-vehicle information systems (IVIS) such as car radios, hand-held or hands-free mobile phones and navigation systems can lead to driver' s distraction, thus impairing road safety (ISO/DIS 26022, 2007). Driving requires concurrent execution of cognitive, physical, perceptual and motor tasks; IVIS typically involve visual–manual–cognitive demands, the same required for the driving task, and will therefore be very likely to cause interference (Wickens, 2002). Such interference is expected to be reflected in an increase of mental workload, conceptually defined as the amount of mental work or effort that an individual makes to perform a task (Xie & Salvendy, 2000). Young and Regan (2007) report that a large and rapidly growing body of research has examined the impact of IVIS on driving performance; most researchers focused on mobile phones, and the general finding is that the physical distraction associated with hand-held phones can result in a significant safety hazard. Nonetheless, previous research (Horrey & Wickens, 2006; Matthews, Legg, & Charlton, 2003; Mazzae, Ranney, Watson, & Wightman, 2004) has shown that driving performance is affected even by IVIS, which do not require visual attention or manual response, such as hands-free phones.

---

[*] Corresponding author. Address: Università di Torino, Dipartimento di Psicologia, Via Verdi 10, 10124 Torino, Italy. Tel.: +39 347 06 43 581.
E-mail addresses: marco.pedrotti@unito.it, marco_pedrotti@libero.it (M. Pedrotti).

1369-8478/$ - see front matter
doi:10.1016/j.trf.2010.12.001

Research on driver distraction has been conducted with several methods and techniques, and there is a general difficulty in comparing results from different studies (Young & Regan, 2007). The International Organization for Standardization (ISO) is making efforts in this sense, and a working standard is under development (ISO/DIS 26022, 2007) for providing a simple and reliable method – the Lane Change Test (LCT) – for quantitatively estimating secondary task demand in a driving context.

Previous research reported a significant correlation between performance and subjective workload measures in the LCT (Wynn & Richardson, 2008). The issue of establishing which methods and measurement techniques are most sensitive to the differential effects of IVIS on drivers' distraction is still an open question (Young & Regan, 2007). Matthews, Davies, Westerman, and Stammers (2000) identified five criteria for selecting a workload index, such as sensitivity (the index should be sensitive to changes in demands for resources), diagnosticity (the index should discriminate sources of demand), selectivity (the index should not be sensitive to extraneous factors – e.g. emotions), obtrusiveness (assessment should not disrupt task performance), reliability (the index should be sensitive to changes in demand over time).

In this context, our research aims at defining an appropriate indicator for driver visual workload, under the assumption that vision is the single most important source of information for the driver (Lansdown, 2000). According to Wickens' (2002) multiple resources model, human vision processes rely on a general pool of mental resources, from which all other processing codes, stages and modalities draw. By adopting such a model, we consider visual workload as a part of mental workload. Purely behavioural measures fail to provide fully satisfactory indexes of these aspects of cognition, hence the need for psychophysiological measures. Within such measures, eye movements have the advantage of being unobtrusive, since they could be collected with remote eyetrackers.

We propose the introduction of eye movement metrics, such as eye blinks and pupil dilation within the simulated LCT, assuming that eye movement metrics such as fixation number, fixation duration, saccadic amplitude, saccadic speed and gaze position would be obviously affected by the use of a IVIS while driving, because of the need to continuously shift gaze from the IVIS to the road and vice versa. While it is well-established that pupil dilation provides a sensible and reliable measure of mental effort (i.e. mental workload) its selectivity and diagnosticity deserve further investigation. Our analysis is focused on eye blink duration as a sensible, reliable and unobtrusive index of visual workload. In this study participants performed a LCT in a driving simulator, in both single- and dual-task conditions. The latter, was manipulated in order to specifically evaluate driver's mental and visual workload due to interaction with an IVIS.

The blink of the eye, the rapid closing and reopening of the eyelid, is believed to be an indicator of both fatigue and workload. It is well known that eye blink rate (BR) is a good indicator of fatigue: results from previous studies (Fukuda, Stern, Brown, & Russo, 2005; Stern, Boyer, & Schroeder, 1994) report that the number of blinks increases as a function of time on task (TOT). Nevertheless, it might be related to aspects of visual (Brookings, Wilson, & Swain, 1996; Veltman & Gaillard, 1996, 1998) and mental (Holland & Tarlow, 1972, 1975; Recarte, Pérez, Conchillo, & Nunes, 2008) workload. Recarte et al. (2008) suggest that the relevant amount of visual attention required by the driving task could lead to a blink inhibition and that the fatigue associated with long driving periods would impair such inhibition. Besides BR, blink duration (BD) has been shown to be affected by visual task demand. In simulated flight tasks, Veltman and Gaillard (1996) and Ahlstrom and Friedman-Berg (2006) report a decrease in BD as visual workload increases. Other studies confirm long-lasting blinks to be a good indicator of drowsiness, whereas very short blinks are often related to sustained attention (Ahlstrom & Friedman-Berg, 2006; Ingre et al., 2006).

The pupil is the part of the iris that allows light to enter the retina. Besides light, the pupil dilates for many reasons such as emotions, loads on working memory and – more in general – because of mental effort. In the latter case, evidence from experimental studies has shown pupil dilation to reflect between and within task variations, as well as between subjects differences (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Goldwater, 1972; Kahneman, 1973). The research conducted by Recarte and Nunes (2000) has shown that performing verbal or spatial-imagery mental tasks while driving has a significant effect on pupil size. Moreover, pupil dilation has been successfully used for distinguishing different levels of difficulty of various cognitive tasks (Beatty & Lucero-Wagoner, 2000).

Notwithstanding, pupil diameter variation represents an innovative measure in the LCT; we analyzed this variable in order to validate the fact that the introduction of an IVIS truly generates measurable workload on the driver, which was further confirmed by the analysis of subjective measures. Once the workload manipulation is validated, we want to show in detail how the visual workload component is indexed by the eye blinks.

In the dual-task conditions, besides higher subjective ratings (Hart, 2006), longer reaction time (Harbluk, Burns, Lochner, & Trbovich, 2007), and lower performance on the IVIS (Wynn & Richardson, 2008), we expect to find larger pupil size (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Goldwater, 1972; Kahneman, 1973; Recarte & Nunes, 2000), and significant differences on pupil size between the different levels of secondary task difficulty we manipulate (Beatty & Lucero-Wagoner, 2000). According to Recarte et al. (2008), we expect a BR inhibition due to the increased visual workload during the dual-task conditions and an impairment of such inhibition as an effect of TOT. According to Veltman and Gaillard (1996), we expect to find a decrease in BD and BR in the dual-task condition. Following Ahlstrom and Friedman-Berg's (2006) findings on BD we expect to find significantly shorter BDs in the dual-task condition.

2. Method

2.1. Participants

Fifteen participants (12 male, mean age = 31, min = 25, max = 36, SD = 4) were recruited and informed about the experiment's general purposes. All of them declared they had valid Italian driving licenses, a minimum of 6 years of driving experience (max = 18, mean = 13, SD = 4), a minimum of 3000 km per year (max = 40,000, mean = 23,000, SD = 10,000) of which, on average, 74% on a familiar path (min = 10, max = 90, SD = 21), and were used to interact with IVIS while driving. Participants were informed about the possibility of giving up (without any consequences) at any time if they did not feel comfortable during the experiment. All of them furnished explicit consent about data recording of their driving performance, eye movements and subjective workload assessment questionnaires. Everyone had normal vision or corrected-to-normal vision (contact lenses, but not glasses were accepted).

2.2. Apparatus

2.2.1. Driving simulation

An *Oktal SCANeR II* driving simulator (Fig. 1) was set up for performing the Lane Change Test (see Section 2.3.3) according to the specifications of the ISO 26022 (ISO/DIS 26022, 2007). The driving simulator software records the vehicle position and dynamics at a frequency of 20 Hz: data logs are saved as *txt* files for offline analysis.

2.2.2. Eye-tracking

An *SMI iView X HED* head-mounted monocular eye-tracker was used. A five-point calibration was made for each participant; calibration was further checked with a laser-pointing device before each driving trial. The eye-tracker software records the participant's gaze point and pupil diameter at a frequency of 200 Hz. Pupil diameter is expressed in pixels, as reproduced by the eye-tracker camera on the workstation screen (resolution 1024 × 768). Data logs are saved as *txt* files, which can be analyzed offline. Room lighting was kept constant during all experiment trials.

2.2.3. Secondary task display settings

A 13 × 17 cm touchscreen display (resolution 800 × 600) was mounted on the dashboard to the right of the steering wheel, where IVIS are usually installed.

2.3. Procedure

2.3.1. Information to participants

All participants were previously informed about the purpose of the test, its procedure, equipment and expected duration. A brief explanation of the Lane Change Test was provided before training on the driving simulator.

Fig. 1. Driving simulator: a car cockpit with steering wheel, accelerator–brake–clutch pedals, and gear lever. The scenario is projected on a 244 × 180 cm screen; the mean distance between participant and screen is 274 cm (depending on participant's height and seat style). A touchscreen for the secondary task is mounted to the right of the steering wheel; the distance between participant and touchscreen is approximately 75 cm.

2.3.2. Training

Three training sessions were performed before conducting the experimental trials; participants first trained on the only driving task for at least 2 min, then trained on the IVIS task alone (1 min minimum). Finally, dual-task training (driving and IVIS concurrently) was performed (2 min at least) to ensure the complete understanding of the tasks to be performed.

2.3.3. Primary task: Lane Change Test (LCT)

The Lane Change Test consists of a driving task where the subject is required to perform at least 18 lane changes on a 3 km straight three-lane road. Road signs, appearing every 150 m on both sides of the road (Fig. 2), indicate Lane changes; the vehicle's speed, controlled by the simulation software, is kept at 60 km/h.

The main purpose of this kind of test concerns the quantitative assessment of driving primary task performance degradation while a concurrent secondary task is performed.

Participants were instructed to perform good lane keeping when driving straight and to begin lane changing as soon as they could see the signs, but not before. No instructions were provided about how to prioritize attention between the driving task and the secondary task; participants were explicitly required to perform the dual-task condition to the best of their capability (ISO 26022, 2007).

Particular care was used for emphasizing that the test aimed not to evaluate the participant's skills but rather how in-vehicle multitasking could negatively affect driving performance.

2.3.4. Secondary task: IVIS

The Surrogate Reference Task – SuRT (Mattes, 2003) – was chosen as secondary task, with the aim of evaluating driving performance degradation caused by a generic visual search task rather than a specific IVIS system. The SuRT (like most commercial IVIS) requires visual perception and manual response: such activities, according to Wickens' multiple resources model (Wickens, 2002), require the same mental resources of the driving task and will therefore be more likely to interfere, possibly causing a performance degradation. Thus, we believe the SuRT is a good compromise between the need for a generic distractor task and the need for using ecologically valid distractor tasks rather than artificial tasks (e.g. arithmetic calculation, letter search, etc.) which may over-estimate the negative effects of dual-task driving (Young & Regan, 2007). Moreover, Recarte et al. (2008) suggest that a continuous visual search task is the key for the study of mental workload measures, and previous research suggests that the workload generated by the SURT task may reflect the amount of workload of a real task (Wynn & Richardson, 2008). A two-column SuRT was set up (Fig. 3): participants were required to double-click on the portion (left or right) of the screen where the target circle is located. Two difficulty levels were used (Fig. 3): an *easy* one with fewer distractors (small circles), and a *difficult* one with more distractors. In both difficulty levels, targets (large circles) have a diameter of 1.4 cm (distractors 0.7 cm).

2.4. Experimental design

A within-subjects design (Fig. 4) was used: two single-task runs were collected, one at the beginning (A = baseline) and one at the end (F = control). Four dual-task runs (B–E) were performed between the single-task runs, two with an easy SuRT and two with a difficult one (see Section 2.3.4), in a randomized order.

Six different driving scenarios were set up, each one having 18 lane change signs: the scenarios differ from the order and type of sign presented. All participants performed all of the six scenarios, in a randomized order. Each scenario has a length of 3500 m: during the first 500 m participants were asked to start the engine and reach the speed of 60 km/h; the vehicle was limited via software to this set point, then drivers could easily keep it constant by simple flat out. When the 500th meter was reached, a START sign appeared on the left and right sides of the road, indicating that the LCT was to be performed. An END sign at the end of the track indicated that the task had been completed (ISO 26022, 2007).

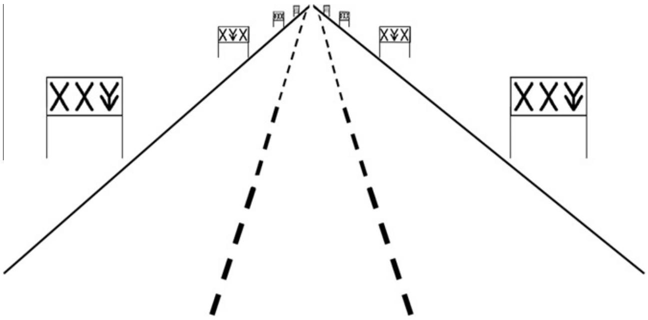
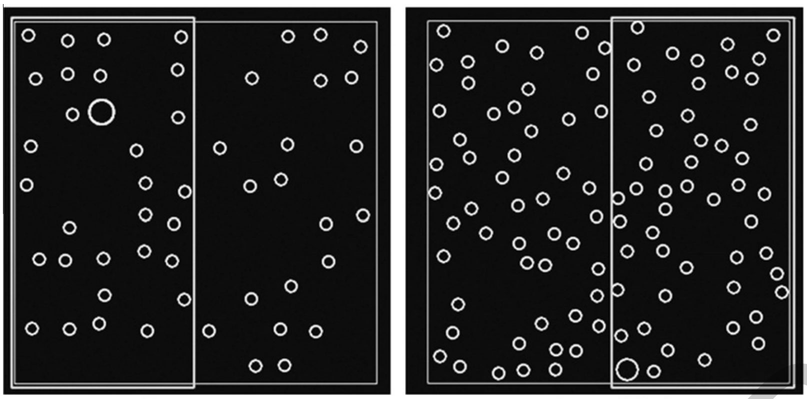


Fig. 2. Road scenario of the LCT (ISO 26022, 2007).



SuRT Easy SuRT Difficult

Fig. 3. Participants has to select as fast as possible the part of the screen where the target circle appears. After clicking on the portion of the screen where the target is located, the system highlights the selected portion with a red rectangle: one further click is required for confirmation, after which a new image is displayed. Twenty images with random localization of the target circle have been created for each of the two difficulty levels.

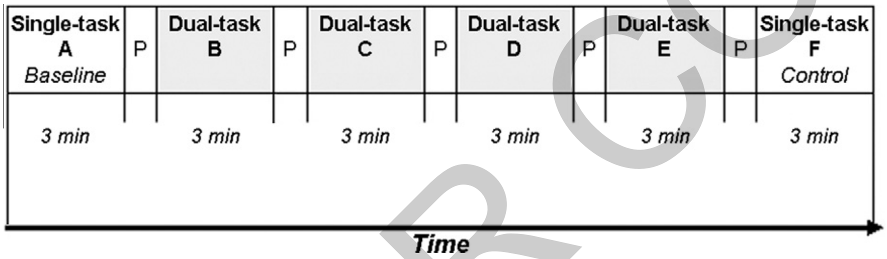


Fig. 4. Experimental design: tasks A and F are the single-task (baseline) runs (lane change task only); tasks B–E are the dual-task runs (lane change task + SuRT task). Pauses (P) between the experimental trials allowed for subjective ratings data collection.

2.4.1. Dependent variables

Mental workload measurement techniques have been classified in two categories: *empirical methods* and *analytical methods* (Xie & Salvendy, 2000), the former gathering data from the operators performing the tasks, the latter being more suitable for prediction purposes in early system design without an “operator-in-the-loop”. According to this taxonomy, we collected dependent variables from three sub-categories of empirical methods, i.e. psychophysiological (BR, BD and Average Pupil Size), subjective (NASA-TLX, RSME) and performance measures (reaction time, IVIS performance). Subjective measures (Hart & Staveland, 1988; Zijlstra, 1993) were collected after each trial.

2.4.1.1. Blink rate. Stern et al. (1994) demonstrated that BR increases as a function of TOT. However, they stated BR could be also affected by other factors, such as perceptual demand of tasks and cognitive variables: experimental evidence was recently provided by Recarte et al. (2008) which came to the conclusion that BR is affected by both mental workload and visual demand, which act in opposition to each other, the former leading to BR increase, the latter to BR decrease.

The event detection software used in the current study selects saccades as primary events using a velocity-based algorithm (e.g. Salvucci & Goldberg, 2000): blinks and fixations are computed and derived from the primary saccade events (SMI – SensoMotoric Instruments, 2009).

2.4.1.2. Blink duration. Stern, Walrath, and Goldstein (1984) reported that little experimental work had been done relating blinks to cognitive activity, and that most studies were restricted largely to BR. The authors focused their literature review on the “endogenous” blink, which they defined as distinguishable from other blinks by the absence of identifiable eliciting stimuli, and they suggested that the endogenous blink could be an indicator of mental processing activity; moreover, they argued that both subject variables and momentary task demands could be reflected by blinks.

Comparing BD across different studies, it is quite difficult and the reasons could be found in the following assumptions: in these studies (a) different tasks are performed and (b) different measurement techniques are used. Concerning (a), we came to the conclusion that BD should be analyzed task by task (e.g. driving, in flight, in air-controlling, etc.) in order to verify whether the kind of task influences BD changes. Concerning (b), we propose that the study of BD – in such applied research contexts – is carried out with video-based remote eye-tracking systems (e.g. Ahlstrom & Friedman-Berg, 2006) because the application of electrooculography (EOG), for the recording of variables of interest (e.g. Veltman & Gaillard, 1996, 1998), is acceptable in laboratory investigations but is, for a variety of reasons, not possible outside laboratories (Stern et al., 1994). While EOG techniques provide more detailed data such as closing portion, reopening portion and closed portion (Stern et al., 1984), they are too invasive, and difficult to adapt to real-world situations. Nonetheless, the use of reliable and standardized eye blink identification algorithms for eyetracking data is needed, in order to make it possible to compare results across different studies.

2.4.1.3. Average Pupil Size (APS). The human pupil primarily dilates as a consequence of luminance (pupillary light reflex) and sensory, mental and emotional events (Beatty & Lucero-Wagoner, 2000). Beatty labelled as TEPRs (Task-Evoked Pupillary Responses) those dilations, which occur as a consequence of cognitive processing and demonstrated that TEPRs can be used as an indicator of processing load (Beatty, 1982).

The eye-tracking system provides pupil diameter (in pixels, see Section 2.2.2) on both the x-axis (*diaX*) and the y-axis (*diaY*): pupil size was computed (blink artifacts were removed) as an ellipse area using the formula:

$$\mathrm{APS} = \mathit{diaX}/2 * \mathit{diaY}/2 * \Pi$$

Subsequently, APS was computed for each of the 90 trials (15 participants × 6 trials), discarding all data points before the START and after the END signs: thus, each APS computed was relative to a period of 3 min of single- (A, F) or dual-task (B–E) run (Fig. 4). One participant data was excluded from the analysis because of poor recording quality.

2.4.1.4. Reaction time (Lane Change Delay). The Lane Change Test can in someway be explained within the probe reaction paradigm (ISO 26022, 2007), with stimuli (the road signs) and responses (initiating lane change manoeuvre). The lane change initiation point is the most significant steering action toward the new lane: the Lane Change Delay (LCD) was calculated as the time difference between the appearance of the lane change sign and the lane change initiation point (ISO 26022, 2007). Thus, we may look at the LCD as a typical reaction time, which is itself a relevant issue from a practical perspective, since driver reaction to events in the visual field (i.e. event detection) is a critical safety issue in the driving task.

2.4.1.5. IVIS performance. Number of correct responses to the secondary task was recorded; given the simplicity of the task, no errors were detected. However, small differences between the *easy* and *difficult* conditions were detected (see Section 3.4).

2.4.1.6. NASA-TLX and RSME scores. NASA-Task Load Index (Hart & Staveland, 1988) ratings were collected after each trial; the NASA-TLX is intended to measure the operator's perceived workload along six dimensions: mental demand, physical demand, effort, own performance, temporal demand, frustration. Since its development, the NASA-TLX has been used in a variety of studies and fields of research, and its reliability and sensitivity have been tested in a consistent number of independent evaluations (Hart, 2006).

Rating Scale for Mental Effort (RSME) is a scale developed by Zijlstra (1993). Perceived mental effort is rated by an indication on a vertical graduated scale. RSME is a unidimensional scale, which strictly investigates mental effort and no other aspects of mental workload. This scale has been widely used in traffic research (De Waard, 1996), since it is a fast and easy method.

3. Results

The mean values for different dependent variables are provided in Table 1. Repeated measures ANOVA (rmANOVA) on the dependent variables was used with a Greenhouse–Geisser correction.

Table 1
Dependent variables analysis results.

Measure	Task					
	A Single-task *Baseline*	B Dual-task	C Dual-task	D Dual-task	E Dual-task	F Single-task *Control*
Blink rate (count)	58 (*66*)	32 (*36*)	38 (*37*)	46 (*57*)	43 (*48*)	70 (*55*)
Blink duration (ms)	152 (*46*)	132 (*47*)	137 (*30*)	128 (*24*)	133 (*33*)	167 (*40*)
Average Pupil Size (px)	2906 (*1133*)	3793 (*1590*)	3804 (*1530*)	3591 (*1339*)	3608 (*1369*)	2706 (*961*)
Reaction time (s)	1.08 (*0.12*)	1.08 (*0.2*)	0.97 (*0.14*)	1.04 (*0.13*)	1.03 (*0.15*)	1.05 (*0.11*)
Reaction time variability (s)	0.24 (*0.07*)	0.46 (*0.12*)	0.42 (*0.09*)	0.43 (*0.19*)	0.42 (*0.1*)	0.25 (*0.1*)
SuRT correct responses (count)	–	77 (*37*)	94 (*30*)	96 (*37*)	105 (*42*)	–
NASA-TLX (0–100)	28 (*14*)	48 (*24*)	41 (*19*)	47 (*23*)	41 (*20*)	25 (*16*)
RSME (0–150)	33 (*21*)	65 (*28*)	62 (*23*)	61 (*22*)	59 (*22*)	27 (*15*)

Note: Means and SD (italic) values recorded on several dependent variables for 15 participants.

3.1. Blink duration

No significant results were obtained in the analysis of BR: high inter-subjects variability (Table 1) may indicate that BR is itself a complex variable, mainly affected by inter-subjects variability and other different factors (see Section 2.4.1.1).

Mean BD values for each experimental trial are reported in Table 1: rmANOVA returned a significant general effect ($F(2, 28) = 4.78$, $p < .05$) across the experimental trials. Since planned contrasts returned no differences within the dual-task conditions, BD values were grouped by averaging values; a significant TOT (*baseline* vs. *control*) effect was observed ($F(1, 14) = 5.9$, $p < .05$), but no significant differences were found between the single- and dual-task conditions. Thus, mean BD only revealed the TOT effect.

Besides considering mean BD, we divided blink events in 10 ms duration classes and analyzed the distribution of BD in the *baseline* condition: a Gaussian-like curve (*skewness* = 0.82; *kurtosis* = −0.32) was obtained (Fig. 5).

Afterwards, BD distribution was analyzed for the *dual-task* conditions: in such conditions, the distribution shifts towards shorter blinks (*skewness* = 1.22; *kurtosis* = 2.15), i.e. more short blinks did occur in *dual-task* conditions.

Finally, BD distribution was plotted for the second single-task condition (*control*): the effect of TOT changes the distribution (*skewness* = 0.1; *kurtosis* = −1.36), which becomes more symmetric, shifting towards longer blinks in respect to the distribution in *baseline*.

With the aim of defining BD for the LCT, we ran a clustering analysis (*k-means*) grouping on the BD distribution in *baseline* (Fig. 5) and found three categories which we labelled *short blinks* (71–100 ms), *medium blinks* (101–170 ms) and *long blinks* (171–300 ms); blinks lasting more than 300 ms were excluded because of poor occurrence.

rmANOVA were carried out within the three BD categories; results returned significant general effect within the *short* ($F(2, 28) = 5.15$, $p < .05$) and *long* ($F(2, 28) = 5.74$, $p < .05$) categories, no effects in the *medium* category.

Planned contrast *baseline* vs. *dual-task* revealed that *short* blinks significantly increase in the visual workload conditions ($F(1, 14) = 6.63$, $p < .05$). In respect to *long* blinks, contrast *baseline* vs. *control* was significant ($F(1, 14) = 6.59$, $p < .05$), revealing the TOT effect. However, we report that the contrast *dual-task* vs. *control* is significant ($F(1, 14) = 5.93$, $p < .05$), which corroborates the finding that TOT has a strong effect on *long* blinks. Total blink count graph (Fig. 6) summarizes the effects described.

3.2. Average Pupil Size (APS)

Since planned contrasts revealed no significant differences within the dual-task conditions we averaged these trials' values; results from the rmANOVA showed a significant general effect ($F(2, 28) = 33.27$, $p < .001$). Planned contrasts showed significant differences between all single- and dual-task experimental conditions: *baseline* vs. *dual-task* ($F(1, 14) = 51.9$, $p < .001$) and *control* vs. *dual-task* ($F(1, 14) = 34.39$, $p < .001$). Although the difference between *baseline* and *control* is not significant, the tendency towards statistical significance cannot be ignored ($F(1, 14) = 4.45$, $p = .053$).

According to our predictions, we wanted to see whether APS could reveal differences between the *easy* and the *difficult* conditions of the SuRT and found a significant general effect ($F(3, 42) = 3.74$, $p = .05$) within the dual-task conditions grouped by difficulty level. We then expected to find differences between the *easy* and *difficult* conditions: planned contrast did not support this thought, revealing instead significant differences ($F(1, 14) = 4.66$, $p < .05$) between the moment of occurrence of these conditions. When the SuRT was presented for the first time (either *easy* or *difficult*), APS was higher than the second time, which might be related to task novelty (Fig. 7).

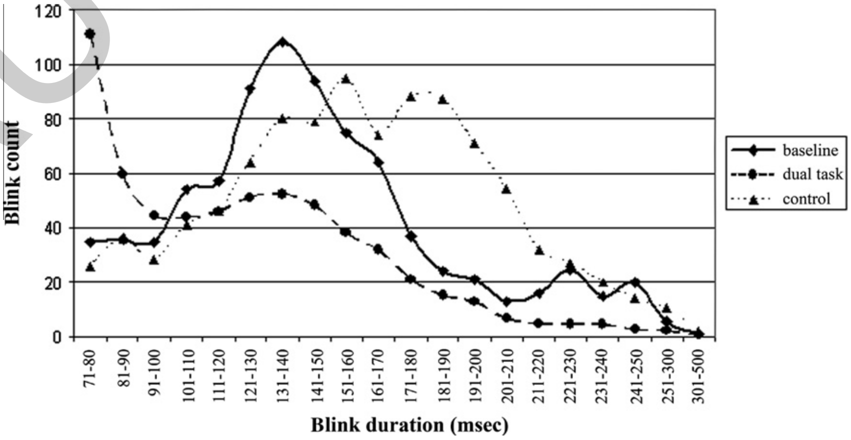


Fig. 5. Distribution of BD in *baseline*, *dual-task* and *control*.

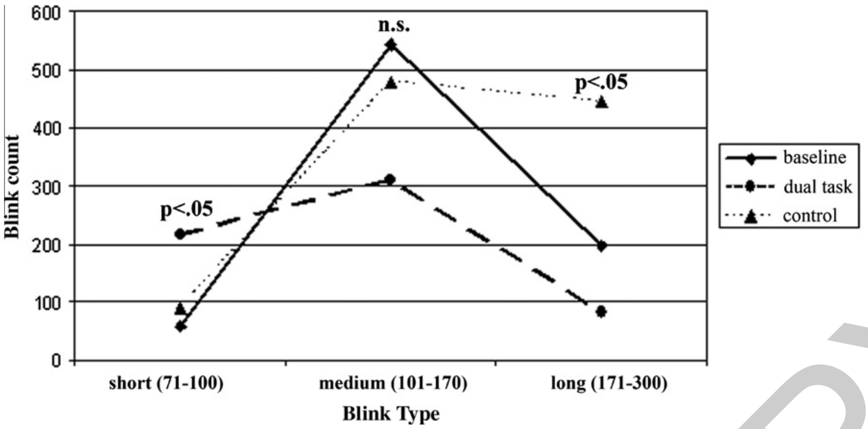


Fig. 6. Total blinks for 15 participants divided in three classes.

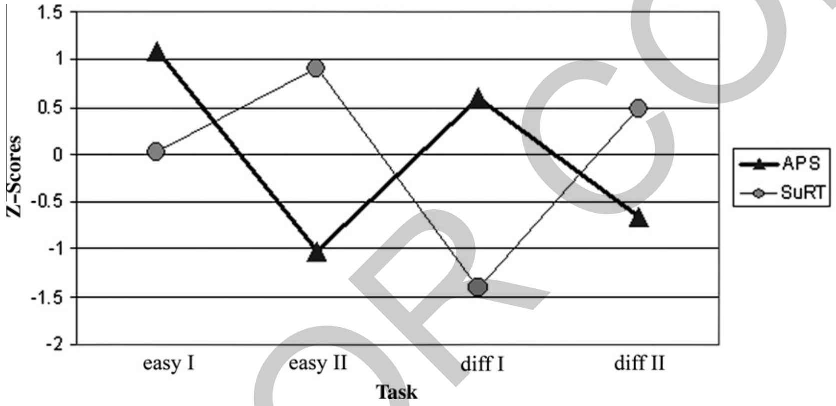


Fig. 7. APS and SuRT responses (*z*-scores) for 15 participants.

3.3. Reaction time (Lane Change Delay)

LCD analysis returned significant results; the rmANOVA returned a significant general effect ($F(5, 70) = 3.15$, $p < .05$) across the six experimental trials, however the effect is attributable to the only C trial, where the LCD is lower, compared to all the other trials (all planned contrasts are significant). The increase in reaction performance in task C comes with both the lowest workload perception within the dual-task trials (see Section 3.5) and the highest pupil dilation values (see Section 3.2).

Since LCD did not vary systematically as a function of the dual-task condition, we tested whether Reaction Time Variability (RTV) was affected by such condition. RTV was calculated for each participant as the standard deviation of LCD on 18 lane changes within each experimental trial. The rmANOVA returned a significant effect ($F(5, 70) = 20.97$, $p < .001$) with all planned contrast between single- and dual-task conditions statistically significant; being significantly sensitive to dual-task conditions, RTV seems a more appropriate indicator of driver distraction than LCD. RTV is higher in the dual-task condition, due to the need to keep the eyes off the road when interacting with the IVIS. This result may have practical implications in the online monitoring of driver attention.

3.4. IVIS performance

Results show an increasing number of responses over time (Table 1), possibly indicating a task habituation effect, which the rmANOVA revealed as statistically significant ($F(3, 42) = 10.4$, $p < .001$). We tested whether the *easy* and *difficult* conditions would show an effect on the IVIS performance and found significant results ($F(3, 42) = 5.20$, $p < .01$); given that the IVIS responses increased between the first and the second time each condition was performed, we carried out the contrast [*easy I, diff I*] vs. [*easy II, diff II*] and found a significant effect ($F(1, 14) = 7.7$, $p > .05$) which we interpreted as task habituation. Fig. 7 shows that lower APS values did occur in the *easy II* and *diff II* conditions, wherein higher SuRT responses were recorded: task habituation leads to more responses with less mental effort.

3.5. NASA-TLX and RSME scores

rmANOVA indicates a significant general effect ($F(5, 70) = 16.72$, $p < .001$), with participants declaring higher perceived workload within the dual-task trials. RSME scores were also collected, and significant positive correlations were found between the two scales in all trials except C (A: $r = .58$, $p < .05$; B: $r = .78$, $p < .01$; C: not significant; D: $r = .82$, $p < .001$; E: $r = .69$, $p < .01$; F: $r = .61$, $p < .05$).

4. Discussion

In respect to previous research, the main findings of this study concern BD. It should be noted that most previous studies did not report exact BD values (means with SD), so that the reader could have an insight into inter-subjects variability of BD. Veltman and Gaillard (1996) found BD to change as a function of workload in a simulated flight task, with shorter BD in difficult with respect to easy conditions. In another simulated flight task, Veltman and Gaillard (1998) found contrasting results, with longer BD in the presumed more difficult task in respect to the easier one, even if BD were shorter in the easy task with respect to the *rest* condition. Ahlstrom and Friedman-Berg (2006) investigated BD in a simulated air-control task, and found BD to decrease linearly as a function of workload; furthermore, they found that BD varied as a function of simulation condition, with shorter BD when participants were not provided a weather display in respect to longer when weather display was provided. Problems for comparing BD across these studies concern different types of both task and measurement techniques (see Section 2.4.1.2): in general terms, a BD decrease seems to be associated with visual workload, and a BD increase may result from TOT. As Recarte et al. (2008) suggested, an inhibition of BR may occur in tasks where visual perception is essential, in such a way that one may lose less relevant information as possible: as time passes (TOT) the mental resources for such inhibition decline, and blink activity starts increasing again.

Our results suggest that besides BR, BD follows this schema, and that BD is a sensitive indicator of visual workload, given the significantly higher presence of short blinks under high visual load conditions (IVIS interaction). Moreover, the possibility of discriminating workload sources – the TOT effect (using long blinks) from the visual load effect – suggests that BD has good diagnosticity. BD reliability has not been tested directly, which would imply a different experimental design with both single- and dual-task within the same eye movements recording trial. Concerning selectivity, we cannot state – with respect to the results of the present study – whether or not BD satisfies this criterion. Finally, unobtrusiveness can be guaranteed if BD is recorded with remote eyetracking techniques: we used a head-mounted system, which is not completely remote but allows-at least-free head movements.

Table 1 shows that BR (like most physiological variables) has a great variability, with SD values even greater than – or very close to – mean values. Surprisingly, BD showed lower variability, and we point out that this result was obtained with 15 participants: further research is needed, with larger sample-size, to validate these findings. The BD classes we introduced seem to provide important information for BD interpretation, in that the queues of the distribution (*short* and *long* blinks) contain relevant information, which is not provided from the centre of the distribution (*medium* blinks), even if more cases – i.e. more blinks – occur within such centre. We even emphasize that our taxonomy of BD (see Section 3.1) needs to be considered and validated within a simple driving task (like the LCT) and with camera-based eye-tracking systems, and does not pretend to apply to other tasks or different blink detection methods.

APS clearly reflected the effort of performing dual-task driving, and an interesting habituation effect across the *easy* and *difficult* conditions of the IVIS task was found, in that APS was significantly higher each time the IVIS condition was performed for the first time (Fig. 7). This suggests that the impact of introducing IVIS tasks in a driving context rapidly leads to lower mental effort as people gain familiarity with such secondary tasks. However, lower mental effort does not really mean that general driving safety is not being impaired by the IVIS.

LCD did not vary as a function of visual workload; one might have expected higher reaction time to the driving task in a dual-task condition, which did not occur. This could be interpreted as subjects investing more mental resources – i.e. trying harder – in workload conditions, which is confirmed by APS values. However, RTV reflected the impossibility to react as fast as possible to *all* the 18 events of the driving task while performing an IVIS task: this suggests that IVIS impair driver attention in a mostly unpredictable manner, which derives from the strategies of dual-task management adopted by each single person. High RTV could be used as an indicator of the level of attention allocated to the IVIS task.

The subjective measures further validated the workload imposed by the IVIS task: the positive correlation between the NASA-TLX and the RSME suggest that both scales may be used as a substitute for each other.

Relevance of this study for system design in automotive mainly concerns the findings on BD, which could be used for online monitoring of driver visual workload, even if further applied research is needed to know whether or not BD could be used outside laboratories. In particular, the relation between BD and real world luminance, air quality and other intervening variables should be investigated.

5. Conclusions

Our research focused on the analysis of psychophysiological variables, which may indicate visual workload; with respect to previous studies, we introduced a detailed analysis of BD. This eye movement measure may have practical implications for the online detection of driver visual distraction, which is the key issue for accident prevention, since vision has been ranked as the single most important source of information for the driver (Lansdown, 2000). BD showed a Gaussian-like distribution in single-task conditions (i.e. only driving task), while its distribution shifted to the left of the curve in dual-task conditions, that is, blink length inhibition may occur to avoid visual information loss.

Future work will regard the validation of BD classes with larger samples, and the analysis of deviation from the normative model as an indicator of driving performance (ISO 26022, 2007) to see whether eye movement metrics may predict such an indicator.

References

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics, 36*, 623–636.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and structure of processing resources. *Psychological Bulletin, 91*, 276–292.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge: Cambridge University Press.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*, 361–377.

De Waard, D. (1996). *The measurement of drivers' mental workload. PhD thesis*, University of Groningen. Haren, The Netherlands: University of Groningen, Traffic Research Centre.

Fukuda, K., Stern, J. A., Brown, T. B., & Russo, M. B. (2005). Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, Space and Environmental Medicine, 76*(7), 75–85.

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin, 77*, 340–355.

Harbluk, J. L., Burns, P. C., Lochner, M., & Trbovich, P. L. (2007). Using the lane-change test (LCT) to assess distraction: Tests of visual–manual and speech-based operation of navigation system interfaces. In *Proceedings of the 4th international driving symposium on human factors in driver assessment, training, and vehicle design* (pp. 16–22), Washington.

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society 50th annual meeting* (pp. 904–908), Santa Monica, CA.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 239–250). Amsterdam: North Holland Press.

Holland, M. K., & Tarlow, G. (1972). Blinking and mental load. *Psychological Reports, 31*, 119–127.

Holland, M. K., & Tarlow, G. (1975). Blinking and thinking. *Perceptual and Motor Skills, 41*, 403–406.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48*(1), 196–205.

Ingre, M., Akerstedt, T., Peters, B., Anund, A., Kecklund, G., & Pickles, A. (2006). Subjective sleepiness and accident risk avoiding the ecological fallacy. *Journal of Sleep Research, 15*(2), 142–148.

ISO/DIS 26022 (2007). Road vehicles – Ergonomic aspects of transport information and control systems – Simulated lane change test to assess in-vehicle secondary task demand. ISO/TC 22/SC 13 - ICS 13.180; 43.040.15.

Kahneman, D. (1973). *Attention and effort*. Engiewood Cliffs: Prentice Hall.

Lansdown, T. C. (2000). Causes, measures, and effects of driver visual workload. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 351–369). Mahwah: Lawrence Erlbaum.

Mattes, S. (2003). The lane change task as a tool for driver distraction evaluation. In H. Strasser, H. Rausch, & H. Bubb (Eds.), *Quality of work and products in enterprises of the future* (pp. 57–60). Stuttgart: Ergonomia Verlag.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance: Cognition, stress and individual differences*. London: Psychology Press.

Matthews, R., Legg, S., & Charlton, S. (2003). The effect of cell phone type on drivers' subjective workload during concurrent driving and conversing. *Accident Analysis and Prevention, 35*, 441–450.

Mazzae, E. N., Ranney, T. A., Watson, G. S., & Wightman, J. A. (2004). Hand-held or hands-free? The effects of wireless phone interface type on phone task performance and driver preference. In *Proceedings of the human factors and ergonomics society 48th annual meeting* (pp. 2218–2222), New Orleans.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery task on eye fixations while driving. *Journal of Experimental Psychology: Applied, 6*, 31–43.

Recarte, M. A., Pérez, E., Conchillo, A., & Nunes, L. M. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology, 11*(2), 374–385.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In A. T. Duchowsky (Ed.), *Proceedings of the eye tracking research and applications symposium* (pp. 71–78). New York: ACM Press.

SMI – SensoMotoric Instruments (2009). *Begaze 2.2 manual*. Germany: Teltow.

Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology, 21*(1), 22–33.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors, 36*, 285–297.

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology, 42*, 323–342.

Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics, 41*(5), 656–669.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177.

Wynn, T., Richardson, J. H. (2008). Comparison of subjective workload ratings and performance measures of a reference IVIS task. In *Proceedings of European conference on human centred design for intelligent transport systems* (pp. 79–88), Lyon.

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. *Work and Stress, 14*(1), 74–99.

Young, K., & Regan, M. (2007). Driver distraction: A review of the literature. In I. J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter, & J. D. Irwin (Eds.), *Distracted driving* (pp. 379–405). Sydney, NSW: Australasian College of Road Safety.

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour. Unpublished doctoral dissertation*. Delft, The Netherlands: Technical University.