www.ietdl.org

Published in IET Intelligent Transport Systems
Received on 10th January 2012
Revised on 18th April 2012
doi: 10.1049/iet-its.2012.0008

ISSN 1751-956X

Can drivers' eye movements be used to monitor their performance? A case study

Y. Yang[1,2] *M. McDonald*[2,3] *P. Zheng*[2,3]

[1]*State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, Southeast University, Nanjing 210096, People's Republic of China*
[2]*Transportation Research Group, Faculty of Engineering and the Environment, University of Southampton, SO17 1BJ, UK*
[3]*Faculty of Maritime and Transportation, Ningbo University, Ningbo 315211, People's Republic of China*
E-mail: yyan1024@hotmail.com

Abstract: This study investigates the relationship between drivers' eye movements and their driving performance, when drivers interact with in-vehicle systems. The impact of a set of in-vehicle visual tasks (VT) on drivers' workload, task performance and eye movements was studied in an on-road experiment with 41 subjects (29 male and 12 female), using an instrumented vehicle. The degree of distraction was differentiated into three groups by cluster analysis based on two sets of selected drivers' eye movement parameters. The results showed that the in-vehicle VT had a negative effect on driving performance, accompanied by a change in drivers' visual behaviour, including a reduction in speed, a higher speed deviation and less predictable steering behaviour. At the same time, a decrease of looking ahead was replaced by viewing the touch screen, and a significant increase in drivers' saccade was observed as the task difficulty level increased. The groups that were the most significantly affected by VT demonstrated a significantly worse driving performance. It is therefore concluded that the deteriorated performance caused by distraction can be diagnosed by visual behaviour change, and it is possible to use eye movements to monitor the change in driving performance.

1 Introduction

The overload of information to drivers has been a topic for traffic safety research since the late 1920s [1], when the first radio receivers were introduced to cars. In modern times, mobile phone usage has increased steadily, and the extra workload and distraction caused by these devices while driving have caused wide-scale research, which has led to a ban on the use of hand-held telephones for drivers in many countries [2]. For example, in the UK study [3], drivers' performance while using a mobile phone was shown to be worse than even under the influence of 80 mg/100 mL of alcohol, the legal limit for the UK. More recently, the increasing use of in-vehicle information systems (IVISs) and advanced driver assistant systems may cause an even greater impact than mobile phones, because of the more intensive information they provide, especially in higher visual demands, and the more complicated driver actions required. Although these systems provide assistance and support to drivers, the higher task demand or distraction caused by these extra stimuli presents an increasing challenge for drivers, as the associated higher workload can adversely affect driving performance, and could lead to information overload, and eventually cause driving errors [4]. According to the US National Highway Traffic Safety Administration, such inattention or distraction, contributes to 25% of all police-reported accidents [5, 6], and this concern is supported by similar comments from the Department for Transport in the UK [7].

A significant amount of research has been conducted on the impact of the increased workload from using in-vehicle systems on driving behaviour and performance [4, 8–14]. For example, the HASTE (Human Machine Interface And the Safety of Traffic in Europe) project used a set of surrogate IVIS tasks to investigate the influence of task load on driver behaviour [9, 10, 12], and similar studies have also been carried out to investigate in-vehicle navigation devices [15, 16], electronic maps [17], traveller information system [18], e-mail [19] and on-board entertainment systems [20]. These and other research related the safety issues in the design of IVISs, including the location of controls, the screens size and brightness, the menus and dialogues between the user and the system, and the channel for information exchange, that is, auditory, visual or haptic. As a consequence, driver interface standards were developed and associated safety issues can be found in Green's review [14].

To enhance road safety and prevent drivers from over loading, various in-vehicle driver monitoring systems have been proposed [21, 22]. However, these distraction monitoring systems are currently based on either measurements of vehicle control, for example, in lane keeping, speed and headway control (e.g. as developed by Volvo and Mercedes-Benz) [23, 24] or drivers' eye and head movements (e.g. by Saab and Toyota) [25, 26], but not both. Research has shown that observing operator's eye and head movements can be used as a non-intrusive tool to understand how the mind acquires and processes visual information [27, 28]. For example, in tennis playing, the eyes always move to the bouncing point one second before the ball arrives at that point [29]. In driving, where we look is guided by where the crucial road and traffic information is. For example, on a curvature road segment, drivers' visual attention is paid mostly on the 'far' and 'near' points (i.e. approaching 'tangent point' and the edge of the lane, respectively) [30, 31]. These findings all suggest that eye movement can provide information on how task activities are organised and how the related data are visually collected by drivers; and they reflected people's strategies for dealing with the primary driving task. In addition, it has been suggested that, during driving, looking away from the road ahead for any period of time was detrimental [32] and many research on eye movement have shown this to be an additional physiological and performance measurement while performing IVIS tasks [13, 33–36]. The visual demand imposed by IVIS tasks can be quantified by drivers' eye movement measurements, and the ISO 15002-1/15002-2 specify guidelines for analysing drivers' visual behaviour with respect to in-vehicle system design based on glance frequency and mean duration [37, 38].

Although many researchers have found IVIS's distractions influences both drivers' visual behaviour and driving performance, limited effort has been made to understand the potential relationship between eye movements and driving performance measurements, which can provide an insight for improving the algorithms used in distraction detection systems. Whereas Kircher and Ahlstrom [39] investigated the accuracy of using a set of vehicle control/driving performance indicators to predict visual distraction, and predicted this accurately to 76%, the severity of the distractions was not considered in the study, that is, the deterioration of driving performance was not considered in the context of different levels of distractions. Drivers may also apply different coping strategies, by engaging less in the IVIS tasks, or making more effort to compensate [40], for example, by prioritising tasks according to their importance [41] or to adopt less demanding strategies, for example, to 'chunk' IVIS tasks into smaller pieces, which can complicate the relationship of visual behaviour and driving performance. This study therefore investigated the effect of IVIS secondary tasks at three levels of difficulty, and explored the relationship between driving performance and eye movements in real-road driving conditions.

2 Method

An on-road experiment was designed to observe drivers' workload, performance and visual behaviour. The experiment induced extra workload on drivers by using a set of visual tasks (VT) to simulate the effects of IVISs, to determine any driving performance change, for example, in vehicle manoeuvring, as well as their eye movement or visual behaviour. VT were chosen because driving is a visually-intensive task, and the overlap of resources, which are needed for simultaneously driving and performing secondary tasks is relatively high, that is, both tasks require visual information processing and manual manipulating, and therefore the impact of the VT on driving performance is more noticeable. The VT were presented to drivers through a touch-screen (TS) mounted in the instrumented vehicle (IV) (see Fig. 1), and were designed to induce an extra workload, where the task demand was not so high as to cause accidents, and for safety reasons, these tasks were tried and tested in pilot runs prior to the full experiment.

During the experiment, drivers were asked to rate subjectively the workload experienced during each task, compared to a previously performed reference task of 'turning on the radio', which is commonly found in IVISs. The secondary tasks were designed to control the 'dose' of the visual distraction at three different difficulty levels, as well as the task demand and anticipated workload. Drivers' performance and visual behaviour at the three difficulty levels were referenced to baseline (BL) 'free' driving, that is, without performing any secondary tasks, to act as a control. During the experiment, each level of the VT was performed three times and the order of task difficulty levels was counterbalanced between drivers. Before performing each secondary task, drivers were also given sufficient time to adjust their driving back to 'normal', to eliminate the effect of any previous tasks.

In total, 41 drivers (29 males and 12 females), aged from 21 to 65, and with a minimum of 3 years' driving experience, participated in the experiment.

2.1 Visual tasks

The set-up of the VT used in this experiment is similar to those used in the ADAM project (Advanced Driver Attention Metrics) [42], which has been used and validated by on-road experiments, for example by Rognin *et al.* [43], Wynn and Richardson [44] and Ranney *et al.* [45]. The advantage of this type of circle reference task is that the levels of difficulty can be adjusted by varying the size and number of distracting or background circles – an example is shown in Fig. 1, with three target circles. For this task, serial images of a group of circles, with one (or more) distinctly larger than the others are displayed on a TS mounted in the vehicle, and drivers are asked to distinguish the larger-sized circle(s) on the display and to touch/click on these circles consecutively. After touching all the larger circles required, they then clicked the 'next' button to move onto the next image for this task. A sound prompt was given to inform drivers for the start of each new image.

After clicking on the 'target' circles on one image, they are asked to click on 'next' to move to the next image when they are ready. A total of 10 s were defined as 'time out', that is, even if not all large circles were correctly clicked, the software will move to the next image regardlessly. Owing to the high demand and vehicle vibration in real-traffic, any click on the edge of the large circles was deemed to be valid in task performance. Several desktop tests and on-road

Fig. 1 *Location of TS and example of a VT in level 3*

pilots using a number of subjects were conducted to evaluate the size and number needed for both the target circle and the other 'distraction' circles to determine the appropriate constitution of the three task levels. In task difficulty level 1, there was one target circle to be clicked, which was obviously bigger than the other distraction circles; in level 2, two targets were required among many more distractions, with reduced size differences; and in level 3, three targets were needed among even more distractions, which were now similar sized. An example of a visual task image for difficulty level 3 is shown in Fig. 1.

2.2 Driving scenario

A real-road driving environment was selected over a laboratory or simulator, as on-road traffic conditions are required to observe drivers' behaviour changes for the objectives of this research. It was believed that on-road experiments [10, 46, 47] can reveal the ground truth of driving behaviour with the use of IVIS, and the effects observed under these circumstances are considered to be valid. However, the design of such experiments can be more complex, as driving performance can also be impacted by the road, traffic and weather conditions, as well as the IVIS distraction. The driving scenario for the experiment therefore had to be carefully selected to ensure the effects of this noise were minimised, and the experimental conditions as consistent as possible. After a series of pilot runs, a 'free-driving' scenario based on a straight, primary rural road (A33) between Winchester and Basingstoke in the county of Hampshire in the UK was selected for the experiment. This route section was chosen because it is parallel to a motorway (the M3 motorway), which diverted most of the traffic and therefore created a quiet, single lane driving situation in each direction, which provides a relatively stable driving environment for drivers, and therefore reduces the workload changes caused by other factors. During the experiment, drivers were asked to keep to a consistent speed of 55 miles per hour (88 kph), and data were collected in dry conditions when there was no vehicle ahead, which could interference with their speed choice.

2.3 Equipment

Data were collected using an IV, which is a Fiat Stilo motor car equipped with sensors to measure driver's usage of controls, speed, lane position, distance to adjacent vehicles and other driving behavioural actions [48].

The IV is equipped with a FaceLAB^{TM} Eye Monitoring System, which is used to investigate drivers' visual behaviour. The System consists of two infrared cameras, which track the driver's head and eye movements, and an on-board computer to receive and process different types of eye and head movement data, including head position indicated by horizontal and vertical coordinates, eye fixations, gaze angles, saccade (fast eye movement) and blink, with a frequency of 60 Hz. The System software provides a model representation of the layout inside the vehicle, which allows the experimenter to define different 'objects' (e.g. windscreen, mirrors, dashboard, TS and the general surroundings), which are used to distinguish the area of interest a driver is looking at (see Fig. 2). This figure shows the 'objects layout' in the System display when a driver is facing the windscreen (shown as a red line initiated from the yellow 'head model') and the eyes are looking at the right mirror (shown as a green line from the head).

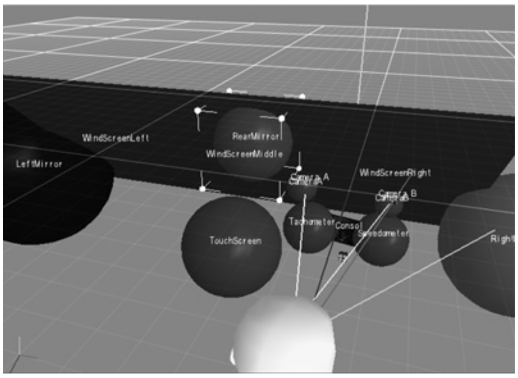

Fig. 2 *Pre-defined gaze objects/areas of interest in the FaceLAB^{TM} system*

An 'operation simulation' system was developed to enable drivers to perform the VT needed for the experiment and to record the drivers' responses, including reaction time (RT) and number of clicking errors, that is, clicks missing or outside the target circles. This system ran on another laptop computer operated by the experimenter, and was linked to the TS and two speakers mounted at the rear of the IV, which sounded to prompt drivers of the start of each image and task.

2.4 Procedure

The content and protocols for this study were reviewed and approved by the Research Ethics Committee and the Research Governance Office at the University of Southampton before the experiment, which was carried out in good weather conditions during August and September 2010, with each driver trial starting at either 09:30 in the morning or 14:00 in the afternoon to avoid the peak traffic hours.

Before each trial, drivers were given instructions for the experiment and a demonstration on how to perform the VT. They were then allowed time to practice on the TS until they were familiar with the manipulation operations required, and to minimise the workload because of the tasks being new. The FaceLAB^{TM} system was also calibrated to account for the face and eye features and dimensions of each driver. Drivers were then instructed to drive normally, without constraint, for half an hour along the M3 motorway to familiarise with the IV. A separate safety observer in the vehicle would check the road and traffic conditions to determine whether it was suitable to perform the VT, and if so, the experimenter would then give instructions for the driver to perform the secondary tasks through the TS, as controlled through the operation simulation system. For each test run, the secondary tasks were presented to the driver three times at each difficulty level, and the order of these levels was balanced among the participants. Drivers were instructed that they should only perform these tasks when it was safe to do so, and they should abort any secondary task if feel unsafe. Drivers were asked to rate the workload immediately after each secondary task, as compared to a control of 'turning on the radio', which had a reference score of 100. For example, they were told a rating of 200 represented a workload which was twice as demanding as turning on the radio.

doi: 10.1049/iet-its.2012.0008

2.5 Definition of performance measurements

Based on recommendation from the literature, the following were used to measure drivers' secondary IVIS task performance, driving performance and eye movements in this study.

'Secondary task performance' was assessed using RT and error percentage (ErP), in which RT is the initialRT, that is, the time period between the moment when each image was presented to drivers and the first correct click of the target circle(s). This variable was calculated as the average RTs for each new image in a visual task. ErP is the percentage of clicks outside the target circle(s), as averaged across each task.

'Driving performance' in free driving was assessed by a combination of longitudinal and lateral vehicle control, as indicated by 'mean speed (MNSP)' during each task, 'speed deviation (SDSP)', 'standard deviation of lane position (SDLP)' and 'steering entropy (SE)'. These have been widely accepted in evaluating driving performance [49–53] and are used to measure any difference in driving performance when conducting in-vehicle secondary tasks. In this study, given that drivers were instructed to maintain a constant speed at 55 miles/h during the experiment, a speed lower than this value can indicate drivers either intentionally compensating for the extra workload by slowing down, or deteriorations in their driving performance. SDSP is another measure of longitudinal control performance, as accident risk increases with more speed variations [54]. Lateral control performance parameters were classified into two types: lane-keeping (or SDLP) and steering behaviour (or SE). SDLP is one of the most commonly used measures to evaluate driving performance for in-vehicle technology, and historically higher SDLP was related to increased accident risk, whereas SE indicates unpredictable steering behaviour, which has proved to be sensitive to steering behavioural changes [55, 56], and the higher this value is, the less predictable the steering. SE was expected to increase as drivers performed the secondary tasks, as they become more distracted. For a detailed description and calculation of SE, see Nakayama's [56].

'Eye movements (or visual behaviour)' were measured by using percentage of time spent looking at the TS and the front or on the road ahead, the per minute blink rate and the saccade rate (Sac Rate). Drivers' eyes typically alternate between saccades and visual fixations, which are the major aspects of visual information extraction. Fixation is relatively 'stationary' eye behaviour, which allows eyes to focus their gaze on the objects being looked at, and to extract this information. Saccade is the eye movement in-between fixations, which leads the fixation from one object to another [29]. The separation and selection of fixation and saccades used in this study were based on temporal criteria, with the former being longer than 140 ms in duration and with a gaze velocity lower than 300°/s; whereas the latter ranged from 10 to 99 ms, in common with other studies [29, 57]. In FaceLAB^{TM}, the blink is a binary signal (true or false), which reports events involving rapid eye closures followed by rapid eye openings [58].

In this study, the 'percentage of time' spent on various specific targets (TS and front) was defined as the percentage of time that subjects spent with their gaze directed to each target region. 'Sac Rate' represents the number of saccades drivers make per minute. Under increased visual workload, drivers made more saccades to shift their visual resource according to the task complexity, and a higher Sac Rate suggested more active searching and therefore greater effort in performing the secondary tasks. Blink rate is the average number of blinks per minute, which is sensitive to both mental and visual demands [59, 60]. Mental workload without visual interference leads to an increase in blink rate, whereas performing VT generally inhibits the blink [59–62].

2.6 Statistical analysis

The statistical analysis of individual variables was computed with SPSS16.0 using ANOVA test. Where the assumptions of ANOVA were violated, for example, the data were skewed (as tested by the Levene test), the robust ANOVA test [63] was conducted. Pairwise tests for significance were carried out with a Bonferroni adjustment [64]. Different subsets of eye movement measurements were used to determine the degree of distraction by K-mean cluster analysis in SPSS. Outliers were identified based on a box plot analysis, and depending on whether they were a 'legitimate' part of the database or not, either a more robust statistical test was used or they were converted to missing values [65].

3 Results

The experiment resulted in 34 complete datasets, including drivers' subjective workload ratings, their secondary task performance from the operations simulation system, their driving performance data from the IV, and their visual behaviour information from the FaceLAB^{TM} system. Data from seven drivers were incomplete due to various technical failures (e.g. poor FaceLAB^{TM} eye tracking, or missing IV data), and were therefore excluded from the analysis. To investigate the correlation of eye movement and driving performance, the remaining 441 cases were divided into different characteristics groups, and the driving performances across these groups were investigated.

3.1 Subjective rating of workload

The results from robust ANOVA showed that drivers' subjective workload rating increased dramatically with the increasing task difficulty levels [Welch, F (2, 187) = 31.610, $p < 0.001$]. Post hoc tests showed that the effect was significant in all three levels, after Bonferrino adjustments ($p < 0.025$). The rated workload increased from a mean of 97.15 in level 1 to 129.38 in level 2 and 167.78 in level 3, as shown in Fig. 3 (the error bars show the 95% confidence interval).

From Fig. 3, it can be seen that the workload at level 1 was comparable to the control task of 'turning on the radio', whereas that at level 3 created a significantly higher workload.

3.2 Secondary task performance

The ANOVA test showed that there were significant increases in drivers' RT [Welch test, F (2, 194) = 3.775, $p < 0.05$] with the increasing task levels, from 2.72 s in level 1 to 2.80 s in level 2 and 3.15 s in level 3, that is, an increased of 2.76% from level 1 to 2, 12.40% from 2 to 3 and 16% overall. Drivers' RT in level 3 was significantly longer than levels 1 and 2, and the difference was more pronounced between levels 3 and 2 than for 2 and 1. No significant effect of ErP across task levels was found. The secondary task performance in each task level is illustrated in Table 1.

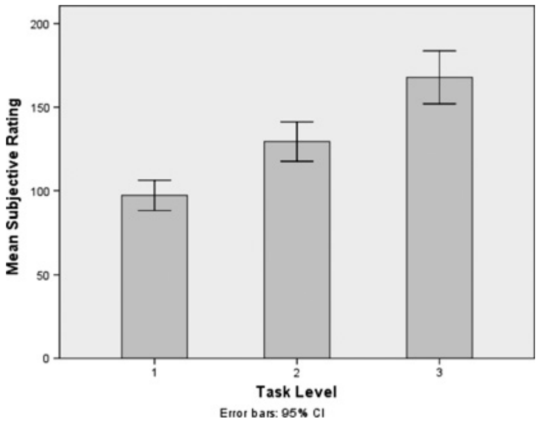

Fig. 3 *Subjective workload rating in the three task difficulty levels*

3.3 Driving performance

In this study, given that drivers were asked to maintain a constant speed at 55 miles/h, SDSP was calculated using the following formula

$$\text{Speed deviation} = \text{sqrt}(\text{sum}((\text{speed} - 55 \times 1.6)^2)/n)$$

where n is the number of the readings in each case.

Driving performance was observed to deteriorate significantly while performing the VT – see Table 2, where 'mean VT' gives the average of all cases when performing the VT, 'Diff BL&VT' shows the difference between BL and mean VT, and 'Diff (%)' represents this percentage, as calculated by

$$\text{Diff} = (\text{mean value in VT} - \text{mean in BL})/\text{mean in BL}$$

There was a significant decrease in MNSP between BL (without secondary task) and levels 1, 2 and 3, with a corresponding increase in SDSP and an increase in SE. The post hoc tests indicate that these differences only exist between BL driving and performing a secondary task, but not significant between the three task levels. A minor increase in SDLP was also observed (from 18.2 cm in BL to 20 cm in level 3), indicating deterioration in lane control, but this effect was not significant.

3.4 Drivers' visual behaviour

Drivers' visual behaviour was dramatically affected by the VT – see Table 3, where 'mean VT' shows the average of all cases whereas performing the VT and 'Diff (%)' indicates the percentage change from BL to performing tasks, as calculated from

$$\text{Diff} = (\text{mean value in VT} - \text{mean in BL})/\text{mean in BL}$$

Table 1 Secondary task performance in each task level

Task level		RT, s	ErP, %
1	mean	2.72	16.35
	SD	1.08	14.91
2	mean	2.80	14.58
	SD	1.08	12.67
3	mean	3.15	14.40
	SD	1.18	13.90

As expected, the percentage of fixation on TS increased significantly as the task difficulty level increased, from a BL of zero (i.e. no need to look at TS in normal driving) to 12.4% in level 1, 13.9% in level 2 and 16% in level 3. Post hoc tests (Student–Newman–Kuels and Tukey's honestly significant difference) showed that this effect was significant at each level. In consequence, the percentage fixation at the front decreased when performing the VT from 77.6% in BL to a mean of 60.4% across all cases whereas performing the tasks. Drivers' Sac Rate also increased from 17.1 to 36.3 times per minute, reflecting more active searching between the TS and front. There was no significant change in blink rate when performing VT. Although some findings were significant, large overlaps were found between each task level, indicating significant individual difference between drivers and the different strategies applied.

3.5 Visual behaviour and driving performance comparisons between different groups

Two sets of eye movement measurements were used to differentiate the degree of distraction, to investigate the correlation of eye movement and driving performance. The 441 cases of drivers' eye movement were divided into three groups by each set of measurements. The first set included two parameters: the percentages of time looking at front and the TS, owing to their importance in detecting when drivers look away, and are expected to be correlated with driving performance. The second set included two further parameters: blink rate and Sac Rate, as they reflect the workload level and active searching behaviour. K-mean cluster analysis was conducted for all the records (in total 441 cases). The results suggested that according to drivers' eye movement, the degree of distraction can be divided into three groups by each set of measurements.

Table 4 showed the result of cluster analysis using two eye movement parameters: the percentage of time looking at front and viewing the TS. Seventy cases (16% of all cases) were classified into group 1, with significantly more time spent looking at the TS (16%) and least time to the front (37%), indicating drivers being seriously distracted. Group 2 comprises the cases where the distracting was moderate, with 12% on the TS and 57% to the front, containing 178 cases (40% of all cases). The remaining 193 cases (44%) were classified as group 3, where the effect of distraction was mild or there was no distraction. The result from the crosstab analysis also showed that the higher the difficulty level, the greater the probability that a subject would be distracted ($p < 0.001$). For example, the group 3 contained all cases for BL driving; whereas most cases in levels 3 (62.8%) shown to be moderately or seriously distracted. The comparison of driving performance between three groups supported this definition. Group 1, that is, the serious distracted group, had a significantly deteriorated performance, with a significantly lower speed and SDSP, indicating the confliction of drivers' mental and manual resource during dual-tasking. This group also showed higher steering adjustment and less predictability in their SE, that is, more and abrupt steering movements. Similarly, compared to the mildly distracted group (group 3), the moderately distracted group (group 2) generally showed worse driving performance; this effect was not statistically significant. However, overlaps between each group appeared to be large.

The result of cluster analysis using the four eye movement parameters, that is, the percentage of time looking at front,

IET Intell. Transp. Syst., 2012, Vol. 6, Iss. 4, pp. 444–452
doi: 10.1049/iet-its.2012.0008

Table 2 Driving performance in each task level

Task level	MNSP[a], kph		SDSP[a], kph		SDLP, cm		SE[a]	
	mean	SD	mean	SD	mean	SD	mean	SD
base line	82.18	4.74	7.29	3.77	18.2	8.29	0.65	0.10
1	78.94	6.28	10.36	5.18	18.6	8.61	0.71	0.07
2	78.17	6.62	10.91	5.91	18.8	7.05	0.72	0.07
3	78.55	5.91	10.53	4.84	20.0	9.08	0.71	0.07
mean VT	78.54	6.17	10.61	5.22	19.2	8.28	0.71	0.08
diff BL&VT	−3.64		3.32		1.0		0.07	
diff, %	−4		46		6		10.52	
df	3,437		3,437		3,437		3,437	
F	8.88		10.22		1.01		17.26	
p	0.00		0.00		0.39		0.00	

[a]Indicates a significant change between level 0 and levels 1, 2 and 3 at the 0.05 level

Table 3 Drivers' visual behaviour in each task level

Task level		TS, %	Front, %	Blink rate, times/min	Sac rate, times/min
base line	mean	0.0	77.6	32.82	17.1
	SD	0.0	19.6	14.91	20.49
1	mean	12.4	63.1	29.04	33.20
	SD	10.4	16.4	13.42	26.00
2	mean	13.9	60.4	29.46	38.1
	SD	10.7	16.6	13.29	32.20
3	mean	16	57.8	28.22	37.4
	SD	13.1	16.3	13.49	28.42
mean VT	mean	14.2	60.4	28.91	36.30
	SD	11.5	18.5	13.76	28.63
df		3,437	3,437	3,437	3,437
F		31.29	25.46	2.066	12.01
p		0.000	0.000	0.104	0.000
Diff, %			−22	−12	113

viewing the TS, and the blink and Sac Rate, is shown in Table 5. Similarly to those in Table 4, Table 5 suggests group 1 was seriously distracted in 35 cases (8% of all cases), with more time spent looking at the TS (16%) and least time to the front (37%). In this group, drivers showed more active searching, indicated by a much higher Sac Rate with 96 times per minute. A higher blink rate also suggested that drivers may have experienced a higher workload. Again, group 2 contained the cases where the distracting was moderate, with 12% on the TS and 57% to the front, containing 132 cases (30% of all cases), with a medium level of searching behaviour and workload, as indicated by the saccade and blink rates. Group 3 seemed to be the least affected by the VT, comprising 274 cases (62%), and the crosstab showed that this group included all BL driving. Again, driving performance among the three groups revealed the same pattern as what was found in the eye movements. The seriously distracted group (group 1)

Table 4 Driving performance compared between different groups of visual behaviour – I

Cases	Group 1		Group 2		Group 3		df	*F*	*p*
Percent	70		178		193				
	15.87		40.36		43.76				
	Mean	SD	Mean	SD	Mean	SD			
TS[a]	33%	9%	11%	7%	5%	6%	2,438	448.92	0.00
front[a]	50%	12%	52%	13%	80%	10%	2,438	330.67	0.00
MNSP, kph [a]	77.47	6.31	79.50	7.08	79.68	5.03	2,438	3.58	0.03
SDSP, kph [a]	11.27	5.72	10.23	5.68	9.22	4.45	2,438	4.47	0.01
SDLP, cm	18.76	8.26	19.01	7.90	18.98	8.66	2,438	0.03	0.98
SE[a]	0.73	0.06	0.70	0.08	0.69	0.09	2,438	5.84	0.00

[a]Difference between the two groups is significant at the 0.05 level (two-tailed)

Table 5 Driving performance compared between different groups of visual behaviour – II

Cases	Group 1		Group 2		Group 3		df	*F*	*p*
Percent	35		132		274				
	7.94		29.93		62.13				
	Mean	SD	Mean	SD	Mean	SD			
TS, %[a]	16%	9%	12%	10%	11%	12%	2,438	3.52	0.03
front, %[a]	37%	14%	57%	14%	71%	17%	2,438	90.19	0.00
blink rate, times[a]	49.99	13.18	29.40	12.23	27.22	12.35	2,438	52.55	0.00
Sac Rate times[a]	96.00	21.71	53.98	13.75	14.09	9.58	2.438	980.48	0.00
MNSP, kph[a]	76.45	4.09	79.38	4.27	79.56	7.03	2,438	4.02	0.02
SDSP, kph[a]	11.96	3.95	9.44	3.93	9.94	5.83	2,438	3.24	0.04
SDLP, cm	0.17	0.06	0.19	0.08	0.19	0.09	2,438	1.06	0.35
SE[a]	0.72	0.07	0.71	0.08	0.69	0.09	2,438	4.12	0.02

[a]Difference between the two groups is significant at the 0.05 level (two-tailed)

showed a worse driving performance, with lower mean speed, higher SDSP and less stable steering adjustments, indicating a higher conflict of drivers' mental and manual resource during dual-taking. This group also showed higher steering adjustment and less predictability in their SE, that is, more and abrupt steering movements. Similarly, compared to the mildly distracted group (group 3), the moderately distracted group (group 2) generally showed worse driving performance; however, this effect was not statistically significant. The difference between groups 2 and 3 is generally not significant, apart from higher steering adjustments observed in group 2. This reflects that this group invested to obtain stable lane control [66] and higher task goal in driving performance, which in turn may induce a higher workload. This explanation is supported by the higher Sac Rate in drivers' eye movements in group 2, compared to group 3. Given that the time spent on viewing the TS between groups 2 and 3 are broadly similar, the less time spent on the front and the higher Sac Rate for group 2 are likely to be owing to the extra attention paid on other driving-related information (such as mirror checking), which contributed to the higher workload.

4 Discussion and conclusion

The performance of in-vehicle VT caused significant deterioration in driving performance, both in terms of longitudinal and in terms of lateral control. Specifically, driving performance deteriorated, with a significant decrease in mean speed, an increase in SDSP and in both minor and large steering wheel adjustments, along with an increase in steering entropy, indicating less predictable SE whereas performing the VTs compared to BL driving. An increase in SDLP of 2 cm was also observed, indicating deteriorated lane control. Although minor, the deterioration in lane keeping is noteworthy, given that the typical increases in SDLP under an intake of alcohol at legal limit range from 2.5 to 4 cm [67]. It is interesting to note that some of the 'worst' performance measurements were found in level 2 instead of level 3. This may be due to drivers being instructed to 'focus on safe driving, and to abandon the secondary task when it is no longer safe'. Drivers may have prioritised their driving over the VT when they found the third level to be too demanding, which in turn 'improved' their driving performance of the most difficult level. As task difficulty level increased, drivers' subjective rating of their workload also increased and their RT to perform the VT increased. As expected, the percentage fixation on the TS increased significantly with each increasing task difficulty level, whereas the time spent on looking to the front decreased. The Sac Rate also increased, which was likely because of the visual transitions between the TS and the areas in front. After classifying all cases of drivers' visual behaviour, it was found that the groups which spent more time on viewing the TS and less time on the front showed a significantly worse driving performance. This is consistent across two sets of measurements for eye movements, particularly, the set in which blink and Sac Rate also reflected the differences in the workload level. 'Therefore drivers' visual behaviour change can be used to detect the deterioration in driving performance'.

These effects on driving performance were observed principally in dual-tasking compared to BL driving, but the variation between task difficulty levels was not statistically significant. As in the validation process of the experiment design, subjects' feedback showed that the discrimination of difficulties between three task levels was significant, and level 3 was already close to the upper limit of human capacity. The non-significant effects in driving performance suggested that the relationship between the workload imposed by secondary IVIS tasks and driving performance was complex and not straightforward, and the variation cannot be explained by looking at the primary driving and secondary task performance alone (or indeed the drivers' subjective workload ratings). It was also possible that the driving performance metrics were not sensitive enough to differentiate among various levels of the secondary tasks. Drivers may have adopted different coping strategies across the three task levels to mitigate the performance deterioration. This was suggested by the saccade behaviour. For example, there was a noticeable increase in Sac Rate when task demand was relatively low, that is, from levels 1 to 2, when fixation on TS increased by 1.5%, and Sac Rate increased from 33 to 38 times per minute. As discussed before, this showed a more active searching behaviour and drivers' effort of balancing driving and secondary task performances by searching between TS and the front. However, when the task became more demanding (from levels 2 to 3), when higher visual searching was needed (fixation on TS increased by 2.1%), drivers did not increase their saccades even further, and rates stayed similar to level 2. Combined with a higher percentage of looking away and

IET Intell. Transp. Syst., 2012, Vol. 6, Iss. 4, pp. 444–452
doi: 10.1049/iet-its.2012.0008

even less time looking to the front, it was likely that drivers looked away from the road ahead, that is, looked at TS for longer durations each time, which in turn caused a higher SDLP. Also, combining the driving performance data with visual behaviour information can reveal extra knowledge in driving behaviour research. For example, a performance trade-off when performing secondary tasks, that is, worse speed control was related to higher percentage of time spent looking at the TS, and more active searching as indicated by saccade behaviour, which suggested that this group had a higher priority on performing secondary tasks. The addition of visual information provides further explanation and an improved understanding of the variability of this relationship. However, given the large overlap between groups, these distractions and effects cannot be determined and explained by eye movement alone.

The relationship between driving performance and eye movements is complicated. However, given that the application of distraction detection monitoring is currently either focused on the detection of driving behaviour, or 'attention away from the road', it is still beneficial to add components of driving performance onto eye/head monitoring-based systems. In a review of existing methods for driver monitoring systems, Dong *et al.* [21] suggested that a 'hybrid' of drivers' visual behaviour and driving performance measurements is a promising method for detecting drivers' attentional status in on-road driving. Such a system can differentiate the source of the distraction by observing where drivers look at, and therefore can process the information of the visual behaviour changes accordingly. By combining this information with driving performance measurements, the system can either provide warning or other assistance information for drivers when a distraction is detected, and to improve safety. Based on the potential of using eye movements to detect driving performance as suggested in this study, future work may be conducted on a larger scale, and in various driving scenarios, to see if it is possible to define the absolute and relative thresholds for visual behaviour and driving performance changes that will enable 'unsafe' levels of distractions to be detected. The general thresholds for the whole population might be difficult to identify because of the overlaps in the data between groups. Thus, while differences maybe statistically significant, care will need to be taken to ensure that the differences are meaningful and robust for practical application. A combination in both general and individual levels might be more suitable.

It is clear that neither driving performance nor eye movement measurements alone do not provide sufficient information on drivers' physiological and attention states or their coping strategies, and therefore cannot detect different types and levels of distractions, which in turn may cause 'false alarms' in detection systems and therefore negatively impact on users' acceptance and adoption of these systems. By monitoring drivers' eye movement 'in addition' to their driving performance, there is the potential to provide improved detections that observe the impact of various levels of driver distractions, which benefit both the design of future in-vehicle systems and help to improve road safety.

5 Acknowledgments

This work was supported in part by the National Science Foundation of China under Grant No. 50678027, and in part by the 111 Project under Grant No. 111-2-05.

6 References

1 Kutila, M.H., Jokela, M., Makinen, T., Viitanen, J., Markkula, G., Victor, T.W.: 'Driver cognitive distraction detection: feature estimation and implementation', *Proc. Inst. Mech. Eng. D, J. Automob. Eng.*, 2007, **221**, pp. 1020–1040

2 Matthews, R., Legg, S., Charlton, S.: 'The effect of cell phone type on drivers subjective workload during concurrent driving and conversing', *Accid. Anal. Prevent.*, 2003, **35**, (4), pp. 451–457

3 Burns, P.C., Parkes, A., Burton, S., Smith, R.K., Burch, D.: 'How dangerous is driving with a mobile phone? Benchmarking the impairment to alcohol' (Transportation Research Laboratory, Crowthorne, UK, 2002), Contract No.: TRL-547

4 Carsten, O.M.J., Nilsson, L.: 'Safety assessment of driver assistance systems', *Eur. J. Transp. Infrastruct. Res.*, 2001, **1**, (3), pp. 225–243

5 'NHTSA driver distraction research: past, present and future'. Available at: http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/233.PDF

6 Wang, J.S., Knipling, R.R., Goodman, M.J.: 'The role of driver inattention in crashes: new statistics from the 1995 Crashworthiness Data System'. 40th Annual Proc. Association for the Advancement of Automotive Medicine, Vancouver, British Columbia, Canada, October 1996, pp. 377–392

7 Department for Transport, Scottish Government, Welsh Assembly Government: 'Reported road casualties Great Britain: 2008 Annual Report'. The Stationery Office, London, 2009

8 Stevens, A., Board, A., Allen, P., Quimby, A.: 'A safety checklist for the assessment of in-vehicle information systems: a user's manual' (Transport Research Laboratory, Crowthorne, UK, 1999), Contract No.: PA3536/99

9 Östlund, J., Nilsson, L., Carsten, O., *et al.*: 'HASTE deliverable 2: HMI and safety-related driver performance' (Institute for Transportation Studies, University of Leeds, 2004)

10 Carsten, O.M.J., Brookhuis, K.A.: 'Issues arising from the HASTE experiments', *Transp. Res. F, Traffic Psychol. Behav.*, 2005, **8**, (2), pp. 191–196

11 Cnossen, F., Brookhuis, K.A., Meijman, T.: 'The effects of in-car information systems on mental workload: a driving simulator study', in Brookhuis, K.A., deWaard, D., Weikert, C.M. (Eds.): 'Simulators and traffic psychology' (Centre for Environmental and Traffic Psychology, Groningen, 1997), pp. 151–163

12 Engström, J., Johansson, E., Östlund, J.: 'Effects of visual and cognitive load in real and simulated motorway driving', *Transp. Res. F, Traffic Psychol. Behav.*, 2005, **8**, (2), pp. 97–120

13 Victor, T.W., Harbluk, J.L., Engstrom, J.A.: 'Sensitivity of eye-movement measures to in-vehicle task difficulty', *Transp. Res. F, Traffic Psychol. Behav.*, 2005, **8**, (2), pp. 167–190

14 Green, P.: 'Driver interface safety and usability standards: an overview', in Regan, M.A., Lee, J.D., Young, K.L. (Eds.): 'Driver distraction: theory, effects, and mitigation' (CRC Press, Boca Raton, Florida, 2008)

15 Walker, J., Alicandri, E., Sedney, C., Roberts, K.: 'In-vehicle navigation devices: effects on the safety of driving performance'. Vehicle Navigation and Information System Conf., Warrendale, PA, 1991, Society of Automotive Engineers, pp. 499–525

16 Ma, R., Kaber, D.B.: 'Effects of in-vehicle navigation assistance and performance on driver trust and vehicle control', *Int. J. Ind. Ergon.*, 2007, **37**, pp. 665–673

17 Tsimhoni, O., Green, P.: 'Visual demand of driving and the execution of display-intensive in-vehicle tasks'. Human Factors and Ergonomics Society 45th Annual Meeting, 2001, pp. 1586–1590

18 Liu, Y.C.: 'Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems', *Ergonomics*, 2001, **44**, (4), pp. 425–442

19 Lee, J.D., Caven, B., Haake, S., Brown, T.L.: 'Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers' attention to the roadway', *Hum. Factors*, 2001, **43**, (4), pp. 631–640

20 Horberry, T., Anderson, J., Regan, M.A., Triggs, T.J., Brown, J.: 'Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance', *Accident Anal. Prevent.*, 2006, **38**, (1), pp. 185–191

21 Dong, Y., Hu, Z., Uchimura, K., Nobuki, M.: 'Driver inattention monitoring systems for intelligent vehicles: a review', *IEEE Trans. Intell. Transp. Syst.*, 2011, **12**, (2), pp. 596–614

22 Liang, Y., Reyes, M.L., Lee, J.D.: 'Real-time detection of driver cognitive distraction using support vector machines', *IEEE Trans. Intell. Transp. Syst.*, 2007, **8**, (2), pp. 340–350

23 Volvo Car Cooperation. Available at: http://www.media.volvocars.com

24 eMercedesBenz. Available at: http://www.emercedesbenz.com

25 Nabo, A.: 'Driver attention – dealing with drowsiness and distraction' (Saab Automobile AB, Trollhattan, Sweden, 2009)

26 Test Driven. Available at: http://www.testdriven.co.uk/lexus-ls-600h/
27 Shinar, D., Gurion, B.: 'Looks are (almost) everything: where drivers look to get information', *Hum. Factors*, 2008, **50**, (3), pp. 380–384
28 Rockwell, R.H.: 'Eye movement analysis of visual information acquisition in driving', *Aust. Road Res. Board Proc.*, 1972, **6**, (3), pp. 316–331
29 Rayner, K.: 'Eye movements in reading and information processing: 20 years of research', *Psychol. Bull.*, 1998, **124**, (3), pp. 372–422
30 Land, M.F., Lee, D.N.: 'Where we look when we steer', *Nature*, 1994, **369**, pp. 742–744
31 Land, M.F., Horwood, J.: 'Which parts of the road guide steering?', *Nature*, 1995, **377**, pp. 339–340
32 Summala, H.: 'Forced peripheral vision driving paradigm: evidence for the hypothesis that car drivers learn to keep in lane with peripheral vision', Vision in Vehicles-VI
33 Hoffman, J.D., Lee, J.D., McGehee, D.V., Macias, M., Gellatly, A.W.: 'Visual sampling of in-vehicle text messages – effects of number of lines, page presentation, and message control'. Eighty-fourth Annual Meeting of the Transportation Research Board, Washington, DC, January 2005
34 Lansdown, T.C.: 'Causes, measures, and effects of driver visual wrokload', in Hancock, P.A., Desmond, P.A. (Eds.): 'Stress, workload, & fatigue' (Lawrence Erlbaum Associates, 2001), pp. 351–369
35 Rognin, L., Grimaud, I., Hoffman, E., Zeghal, K.: 'Assessing the impact of a new instruction on air traffic controller monitoring tasks'. Int. Conf. on Human-Computer Interaction in Aeronautics, Toulouse, France, 2004, AAAI Press
36 Wierwille, W.W.: 'Visual and manual demands of in-car controls and displays', in Peacock, B., Karowowski, W. (Eds.): 'Automotive ergonomics' (Taylor & Francis, London, 1993), pp. 299–320
37 ISO 15007-1: 'Road vehicles –measurement of driver visual behaviour with respect to transport information and control systems. Part 1: definitions and parameters'
38 ISO 15007-2: 'Road vehicles –measurement of driver visual behaviour with respect to transport information and control systems. Part 2: equipment and procedures'
39 Kircher, K., Ahlstrom, C.: 'Predicting visual distraction using driving performance data', *Ann. Adv. Autom. Med.*, 2010, **54**, pp. 333–342
40 Meister, D.: 'Behavioral foundations of system development' (Wiley, New York, 1976)
41 Wickens, C.D.: 'Engineering psychology and human performance' (HarperCollins, New York, 1984, 2nd edn.)
42 Mattes, S.: 'The lane-change-task as a tool for driver distraction evaluation', in Strasser, H., Kluth, K., Rausch, H., Bubb, H. (Eds.): 'Quality of work and products in enterprises of the future' (Ergonomia, Stuttgart, 2003), pp. 57–61
43 Rognin, L., Alidra, S., Val, C., Lescaut, A.: 'Occurrence of secondary tasks and quality of lane changes', in Harris, D. (Ed.): 'Engineering psychology and cognitive ergonomics' (Springer-Verlag, Berlin, Heidelberg, 2007), pp. 397–406
44 Wynn, T., Richardson, J.H.: 'Comparison of subjective workload ratings and performance measures of a reference IVIS task'. The Europe Conf. on Human Centred Design for Intelligent Transport System, Lyon, France, April 2008
45 Ranney, T.A., Scott Baldwin, G.H., Vasko, S.M., Mazzae, E.N.: 'Measuring distraction on potential of operating in-vehicle devices' (National Highway Traffic Safety Administration, Washington, DC, 2009), Contract No.: DOT HS 811 231
46 Brookhuis, K.A., De Waard, D.: 'Why is driver impairment difficult to assess?', in Rothengatter, T., Huguenin, R.D. (Eds.): 'Traffic and transport psychology: theory and applications' (Elsevier, Oxford, 2004), pp. 231–244
47 Brookhuis, K.A., vanWinsum, W., Heijer, T., Duynstee, M.L.: 'Assessing behavioural effects of in-vehicle information systems', *Transp. Hum. Factors*, 1999, **1**, (3), pp. 261–272
48 McDonald, M., Brackston, M.: 'The role of the instrumented vehicle in the collection of data on driver behaviour', *IEE Colloq. Monit. Driver Veh. Perform.*, 1997, **122**, p. 7/1–3
49 Ma, R., Kaber, D.B.: 'Situation awareness and workload in driving while using adaptive cruise control and a cell phone', *Int. J. Ind. Ergon.*, 2005, **35**, (10), pp. 939–953
50 Tornros, J.E.B., Bolling, A.K.: 'Mobile phone use – effects of handheld and handsfree phones on driving performance', *Accid. Anal. Prevent.*, 2005, **37**, (5), pp. 902–909
51 Brookhuis, K.A., De Waard, D.: 'On the assessment of criteria for driver impairment: in search of the golden yardstick for driving performance'. Second Int. Driving Symp. on Human Factors in Driver Assessment, Training and Vehicle Design, Park City, Utah, USA, July 2003
52 Tornros, J.E., Bolling, A.: 'Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments', *Transp. Res. F, Traffic Psychol. Behav.*, 2006, **9**, (4), pp. 298–306
53 Tornros, J.E., Bolling, A.K.: 'Mobile phone use – effects of handheld and handsfree phones on driving performance', *Accid. Anal. Prevent.*, 2005, **37**, (5), pp. 902–909
54 Aarts, L.T., van Schagen, I.N.L.G.: 'Driving speed and the risk of road crashes: a review', *Accid. Anal. Prevent.*, 2006, **38**, (2), pp. 215–224
55 Boer, E.R.: 'Behavioral entropy as an index of workload'. Human Factors and Ergonomics Society Annual Meeting Proc., Proc. 3: Individual Differences, 2000, vol. 4, pp. 125–128
56 Nakayama, O., Futami, T., Nakamura, T., Boer, E.R.: 'Development of a steering entropy method for evaluating driver workload'. Int. Congress and Exposition, Detroit, Michigan, USA, March 1999
57 Di Stasi, L.L., Alvarez-Valbuena, V., Canas, J.J., *et al.*: 'Risk behaviour and mental workload: multimodal assessment techniques applied to motorbike riding simulation', *Transp. Res. F, Traffic Psychol. Behav.*, 2009, **12**, (5), pp. 361–370
58 FaceLAB5 user manual, Seeing machines, Canberra, Australia, 2009
59 Fogarty, C., Stern, J.A.: 'Eye movements and blinks: their relationship to higher cognitive processes', *Int. J. Psychophysiol.*, 1989, **8**, pp. 35–42
60 Recarte, M.A., Perez, E., Conchillo, A., Nunes, L.M.: 'Mental workload and visual impairment: differences between pupil, blink, and subjective rating', *Spanish J. Psychol.*, 2008, **11**, (2), pp. 374–385
61 van Orden, K.F., Limbert, W., Makeig, S., Jung, T.P.: 'Eye activity correlates of workload during a visiospatial memory task', *Hum. Factors*, 2001, **43**, (1), pp. 111–121
62 Wilson, G.F., Fullenkamp, P., Davis, I.: 'Evoked potential, cardiac, blink and respiration measures of workload in air-to-ground missions', *Aviat. Space Environ. Med.*, 1994, **65**, pp. 100–105
63 Buning, H.: 'Robust analysis of variance', *J. Appl. Stat.*, 1997, **24**, (3), pp. 319–332
64 Abdi, H.: 'The Bonferroni and Šidák corrections for multiple comparisons', in Salkind, N. (Ed.): 'Encyclopedia of measurement and statistics' (Sage, Thousand Oaks, 2007)
65 Osborne, J.W., Overbay, A.: 'The power of outliers (and why researchers should always check for them)', *Pract. Assess., Res. Eval.*, 2004, **9**, (6), pp. 1–10
66 MacDonald, W.A., Hoffmann, E.R.: 'Review of relationship between steering wheel reversal rate and driving task demand', *Hum. Factors*, 1980, **22**, pp. 733–739
67 Brookhuis, K.A.: 'How to measure driving ability under the influence of alcohol and drugs, and why', *Hum. Psychopharmacol.*, 1998, **13**, pp. 64–69

doi: 10.1049/iet-its.2012.0008