Detrimental effects of carryover of eye movement behaviour on hazard perception accuracy: Effects of driver experience, difficulty of task, and hazardousness of road

Peter J Hills¹, Catherine Thompson², & J Michael Pake³

¹ – Corresponding author: Psychology Research Group, Bournemouth University, Talbot Campus, Fern Barrow, Poole, Dorset, UK, BH12 5BB

² – School of Health Sciences, University of Salford, Salford, UK, M5 4WT

³ – Department of Psychology, Anglia Ruskin University, East Road, Cambridge, UK, CB1 1PT

Email: phills@bournemouth.ac.uk
Abstract

Novice drivers are more likely to be involved in road accidents than experienced drivers and this relates to their lower performance in hazard perception tasks. Hazard perception performed under dual task conditions is also affected differentially due to driver experience. In this study, we explore the detrimental effect of vertical eye-movement carryover from one task to a second task in drivers of different levels of experience, whilst accounting for road conditions. Participants searched letters presented horizontally, vertically, or in a random array. Following this, they identified a hazard in a driving scene. Carryover of eye movements from the letter search to the driving scene was observed and participants were quicker and more accurate when responding to a hazard following horizontal scanning, compared to following vertical and random scanning. Furthermore, while carryover of eye movements was equivalent for all participants, the negative effect it had on hazard identification accuracy was less prominent in experienced drivers, especially when viewing the most hazardous of images. These results indicate that carryover of eye movements is another potentially distracting effect that can impact on the ability and safety of novice drivers.

Keywords:

Visual search; eye movements; carryover; hazard perception; driver experience; in-car displays
Detrimental effects of carryover of eye movement behaviour on hazard perception accuracy: Effects of driver experience, difficulty of task, and hazardousness of road

Novice drivers are more than twice as likely to be involved in a road-traffic accident than more experienced drivers (Cooper, Pinili, & Chen, 1995). The rate of involvement in accidents decreases quickly over the first two years of driving (Fisher, Pollatsek, & Pradhan, 2006). The most reliable behavioural correlate of crash involvement is hazard perception (Drummond, 2000; Hull & Christie, 1992; McKenna & Horswill, 1999; Pelz & Krupat, 1974; Quimby, Maycock, Carter, Dixon, & Wall, 1986; Transport and Road Research Laboratory, 1979). Hazard perception occurs when a driver detects, evaluates, and reacts to driving situations that have a high likelihood of causing a collision (Crundall, Chapman, Trawley, Collins, Van Loon, Andrews, & Underwood, 2012) and it tends to improve with driver experience. Novice drivers have lower hazard perception accuracy (Crundall, Underwood, & Chapman, 1999; McKenna & Crick, 1991; Renge, 1998; Sagberg & Bjørnskau, 2006) and slower reaction times than experienced drivers (Deery, 2000; Hosking, Liu, & Bayly, 2010; McKenna & Crick, 1991; Quimby & Watts, 1981; Sexton, 2000). These effects hold even when age is controlled for (Maycock & Lockwood, 1993; McKenna & Crick, 1991; Quimby et al., 1986). Though this might depend on the quality of the hazard perception test given failures to find group differences in hazard perception (Chapman & Underwood, 1998; Underwood, 2000). One of the key reasons why novice drivers show poorer hazard perception (and are involved in more accidents) is due to their visual search strategies (McKinght & McKnight, 2003; Pradhan, Hammel, & DeRamus, 2005). Experienced drivers show more horizontal spread of search than novice drivers, therefore allocating attention to more relevant areas of space (Crundall, Chapman, Phelps, & Underwood, 2003; Crundall & Underwood, 1998; Falkmer, & Gregersen, 2005; Konstantopoulos, Chapman, & Crundall, 2010; Mourant & Rockwell, 1972; Underwood, Crundall, & Chapman, 2002, Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003; Wallis & Horswill, 2007). Conversely, novice drivers tend to show more spread of search in the vertical plane relative to experienced drivers (Chapman & Underwood, 1998; Crundall & Underwood, 1998; Evans 1991; Mourant & Rockwell 1972). These eye-movement differences are more pronounced when viewing more hazardous driving scenes: Experienced drivers adapt their eye-movement pattern to suit the hazardousness of the road, whereas novice drivers do not (Chapman, Underwood, & Roberts, 2002; Crundall & Underwood, 1998; Falkmer & Gregersen, 2001; Garay-Vega, Fisher, & Pollatsek, 2009; Lee, Klauer, Olsen, Simons-Morton, Dingus,
Ramsey, & Ouimet, 2008; Underwood, 2007). These results indicate that with experience, drivers learn to anticipate the location of potential road hazards. This experience tends to develop within three years of driving (Crundall et al., 2003).

The pattern of eye movements in experienced drivers indicates that visual search is driven by skill, experience, and familiarity (Ranney, 1994). When viewing natural scenes, in general, observers tend to show highly stereotypical eye movements, focusing on the most informative part of a visual scene (Loftus & Mackworth, 1978; Mackworth & Morandi, 1967) and extracting the most diagnostic information for encoding (Hills & Pake, 2013). Eye movements are also affected by salient objects in the visual scene capturing attention (Buschman & Miller, 2007; Connor, Egeth, & Yantis, 2004). For drivers, these can include hazards generally (Chapman & Underwood, 1998) and specifically their location in the visual field and whether they are moving (Underwood, Chapman, Berger, & Crundall, 2003). In Itti and Koch's (2000) model of attention, top-down and bottom-up factors such as those described can interact to affect attention and eye movements. Indeed, two parallel pathways - focusing on local feature saliency and global contextual features - have been proposed for the control of eye movements and attention (Torralba, Oliva, Castelhano, & Henderson, 2006).

There is another factor that influences eye movements during hazard perception when viewing road scenes. Using relatively realistic driving images and videos, Thompson and Crundall (2011) demonstrated that eye movements from a previous task can carry over to a second task even when the tasks are unrelated. Their participants performed a letter-search task with strings of letters that were arranged horizontally, vertically, or randomly across the screen. Immediately following this, participants saw a road scene (or video clip) and were asked to memorise it (Experiment 1), rate it for hazardousness (Experiment 2), or respond to the onset of the hazard (Experiment 3). Even though the time spent completing the letter search was minimal, the orientation of letters in this task influenced the allocation of attention and eye movements when viewing the road scene, with increased vertical search following the vertically orientated letter-search task (see also Hills, Thompson, Jones, Piech, Painter, & Pake, 2016; Thompson, Howting, & Hills, 2015). In addition, responses to the hazards in Experiment 3 were made significantly quicker following letters presented horizontally compared to letters presented randomly or vertically.

This finding may have some important implications for real world driving scenarios. As the driving task is characterised by a predominantly horizontal spread of search (as described above) secondary displays (in-car and road-side) which induce a more vertical search could influence subsequent (primary) search, and may even be detrimental to the
detection of hazards. Potentially, the presentation of a secondary task may be adding to the
demands of the driving task. It is well known that distraction is the major contributory factor
in crashes (Lestina & Miller, 1994). In an extensive prospective study on distraction, Klauer,
Dingus, Neale, Sudweeks, and Ramsey (2006) found that approximately 78% of crashes and
65% of near-crashes reported were caused by some form of inattention ( ). Distraction is
commonly caused by mobile phones or other in-car displays that involve searching
(conceptually similar to the letter-search task employed by Thompson & Crundall, 2011).
Furthermore, novice drivers are more susceptible to distraction related accidents (Stutts,
Reinfurt, Staplin, & Rodgman, 2001).

Given the role of distraction in crashes, it is important to understand how carryover of
eye movements affects drivers with different levels of experience. Secondary tasks that add to
the mental workload of the driver tend to decrease the driver's scanning (Recarte & Nunes,
would suggest that this is due to the attentional spotlight narrowing to allow focus on the
more central features of the task. This restricts the potential to be distracted by external
information. Experienced drivers have a more proficient search strategy and deploy their
attentional resources more effectively than novice drivers (Crundall, et al., 2003). While we
might assume that this leads them to be less affected by dual tasks, McKenna and Farrand
(1999) have shown that experienced drivers actually suffer greater interference in hazard
perception under dual task conditions than novice drivers. This might be because novice
driver's performance might be too low to be affected by such additional pressures, whereas
there is more potential for performance detriment in experienced drivers. The letter string
task of Thompson and Crundall (2007) intermixed with a driving task can be considered a
dual task procedure, therefore we might expect that experienced drivers will be more affected
by carryover than less experienced drivers. This effect should be enlarged when the primary
task is more difficult than when it is easier as this further adds to the mental workload.

While considering the role of hazard perception, it is important to consider the nature
of the hazards. We have noted that experienced drivers are able to deploy more horizontal
scanning to more hazardous driving scenes than novice drivers. This is based on their
attentional capacity. Indeed, when viewing undemanding driving scenes, experienced drivers
can devote much more of their scanning time to non-road related aspects of the scene
(Hughes & Cole 1986, Green 2002). When there are demands placed on the driver, there are
fewer resources available to allow for this (from perceptual load theory, Lavie, 2010).
Attention will therefore be more limited (Underwood, 2007). Much research has indicated
that hazardousness is the key factor associated with attentional demand in the driving task. However, difficulty of task and hazardousness should not be conflated. Indeed, there is little to no correlation between ratings of hazardousness and hazard perception accuracy (McGowan & Banbury, 2004) or response time (Ferrand & McKenna, 2001). Therefore, in order to assess whether the carryover effect influences participants of differing levels of experience, it is necessary to assess whether the effect varies when the task is more difficult (the hazards are harder to spot) and when the hazards are potentially more hazardous.

The literature is equivocal of the notion that experienced drivers display more horizontal scanning than novice drivers. They also display more horizontal scanning for more hazardous roads than novice drivers. Given that the carryover of eye movements interferes with hazard detection latency (Thompson & Crundall, 2011) and accuracy (Hills et al., 2016), potentially by adding to the task difficulty, it might be that it occurs more for hazardous road scenes than less hazardous ones. This is an empirical question rather than one based on theory. Furthermore, given that with increased task demand, the eye movements of experienced drivers are more affected than novice drivers, we might expect the carryover effect (in eye movements) to be larger for experienced drivers, especially when the hazard perception task is more difficult (not necessarily when it is more hazardous). In order to assess whether the carryover of eye movements effect is the same for drivers of different levels of experience and is the same across road images of different levels of hazardousness and difficulty, we ran a study based on the paradigm of Thompson and Crundall (2011): Participants performed a series of trials in which they completed a letter-search task followed by a hazard identification task.

**Method**

*Participants*

Eighty two (28 male, modal age 22) staff and students from Anglia Ruskin University took part in this study. Participants were paid £3 or given course credits for participation. All participants reported normal or corrected-to-normal vision such that they would be legally able to drive in the UK. Table 1 shows a breakdown of the participant characteristics and whether there were any significant differences across conditions.

*Materials*

Stimuli included were letter strings and 108 road images taken from Thompson and Crundall's (2011) Experiment 1 and from the UK Driving Test. All stimuli were presented onto a 17” (1280 x 1024 pixel) LCD full colour monitor using E-Prime Pro 2. Eye
movements were recorded using a Tobii 1750 eye-tracker (Falls Church, VA), with embedded infrared cameras with a sampling rate of 50Hz. The eye-tracker emits near infrared light, which reflects from a participant’s eyes, and is then detected by the eye-tracker’s camera. The minimum fixation duration was 100ms and the fixation dispersion threshold was 100 pixels based on the default settings for the Tobii eye-tracker.

The letter strings consisted of nine characters. These were presented in black Verdana size 18 font on a white background. Each letter subtended 0.95° x 0.95° of visual angle and was situated within an invisible 9 x 9 grid, subtending 9° x 9° of visual angle. The letter position depended upon the orientation of the search: letters were arranged randomly across the grid in the random search; down the horizontal centre line in the vertical search; and across the vertical centre line in the horizontal search. In every search, there were either 5 consonants and 4 vowels, or 6 consonants and 3 vowels. Letters could be shown in upper or lowercase. The letter ‘I’ was not included as it could have been mistaken for a lower-case ‘l’; participants were made aware of this during the experiment instructions. When the letter string included two of the same vowels, they were counted as two, rather than one vowel.

Stimuli also included 108 road photographs (35.14° x 28.07°) taken from a driver’s perspective in urban, suburban, and rural settings. These images each contained a single hazard (such as a pedestrian stepping onto the road or conflicting traffic). Forty independent drivers rated the images to ensure that they contained only one consistent hazard. These raters viewed each road image and identified the hazard contained within it and also rated the hazardousness of each situation on a scale from 1 (little/no likelihood of causing a collision) to 5 (very high likelihood of causing a collision). The final 108 images were selected on the basis that they contain reliable hazards and the inter-rater reliability was more than .91. The images were grouped into three ‘hazardousness’ categories of low (hazard score of less than 3.5), medium (hazard score of between 3.5 and 4.2), and high hazardousness (hazard score of above 4.2) to allow for an equal number of images in each category. These values were chosen to create an equal number of stimuli in each category of hazardousness. While using still images is not reflective of actual driving, Thompson and Crundall (2011) have shown that a similar pattern of carryover of eye movements is found when using video images as with still images.
Table 1.

Participant details showing the Number of participants in each group, their age (years), number of males, mean length of time holding a licence (months), and mean self-reported experience with driving (out of 7). Group differences were compared. Note that for length of time holding a licence and self-reported experience the comparison was made only between the two groups of drivers.

<table>
<thead>
<tr>
<th></th>
<th>Non-Divers</th>
<th>Novice Drivers</th>
<th>Experienced Drivers</th>
<th>Test of group differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>26</td>
<td>36</td>
<td>20</td>
<td>$\chi^2(2) = 65.33, p &lt; .001$</td>
</tr>
<tr>
<td><strong>Mean Age and range (years)</strong></td>
<td>22.54 (18-48)</td>
<td>22.57 (18-36)</td>
<td>25.00 (18-47)</td>
<td>$F(2, 79) = 1.29, MSE = 35.21, p = .28$</td>
</tr>
<tr>
<td><strong>N male</strong></td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>$\chi^2(2) = 0.33, p = .99$</td>
</tr>
<tr>
<td><strong>Mean (and range) length of time of holding licence (months)</strong></td>
<td>0 (0)</td>
<td>16 (0 - 35)</td>
<td>99 (36 - 370)</td>
<td>$t(54) = 5.45, p = .001$</td>
</tr>
<tr>
<td><strong>Mean (and range) self-reported driving experience (out of 7)</strong></td>
<td>0 (0)</td>
<td>2.58 (1 - 5)</td>
<td>5.56 (4 - 7)</td>
<td>$t(54) = 9.32, p &lt; .001$</td>
</tr>
</tbody>
</table>
**Design**

A mixed 3 x 3 x 3 design was employed. The within-subjects independent variables were the orientation of the letter strings (horizontal, vertical, or random) and the hazardousness of the road (low, medium, and high). The between-subjects variable was driver experience (non-driver, novice driver, and experienced driver) based on number of months since obtaining a licence. Non-drivers were those without a licence and drivers with a licence were separated into two groups according to the finding of Crundall et al. (2003) that experience (in terms of eye movement strategies employed) develops over a three year period. Drivers who had held a licence for less than 36 months were classed as novices and those who had held a licence for 36 months or longer were classed as experienced drivers. In a second analysis to establish whether difficulty of the hazard image affected hazard perception, we replaced the hazardousness variable with difficulty (with the levels: easy, moderately difficult, and hard). Hazard detection difficulty was based on the mean hazard perception accuracy in the control (random letter-string orientation) condition. Eye movements were recorded to the road image and direction of the eye movements was analysed. Accuracy and response time to identify each hazard was measured.

**Procedure**

Once participants gave their informed consent, they were seated with their head in a chinrest 60cm from the screen. Participants eye movements were then calibrated to the eye-tracker. This involved participants following a blue circle move around a screen to nine locations with their eyes. Once successfully calibrated, participants completed the letter search/hazard identification task.

In each trial, participants were presented with the letter search task followed by the hazard identification task. Each letter search began with a fixation cross shown to the centre of the screen for 500ms, participants then viewed strings of letters and were asked to count the number of vowels present (3 or 4) and respond using the keyboard. The letters remained on the screen until a response had been made and feedback was then given (a red or green blank screen) for 1000ms. Accuracy of this vowel counting did not vary with orientation of the letter strings, $F(2, 158) = 1.58, MSE = 0.01, p = .209$, however, reaction time did, $F(2, 158) = 8.72, MSE = 221021, p = .001^1$. Specifically, reaction time to horizontal letter strings was faster than to vertical letter strings (mean difference = 263ms, $p < .001$) and random

---

1 When Mauchley's Test of sphericity was significant for all analyses in this paper, the Huynh-Feldt correction was applied since the epsilon values were above .7. Here we report the uncorrected degrees of freedom, but corrected significance level. The Bonferroni correction was applied to the alpha level of all post hoc analyses throughout the paper.
letter strings (mean difference = 194ms, \( p = .031 \)). There was no difference in reaction times between the random and vertical letter strings (mean difference = 70ms, \( p = 1.00 \)). These effects did not interact with driver experience, \( F(4, 158) = 2.52, MSE = .01, p = .066 \) (accuracy) and \( F(4, 158) = 0.40, MSE = 221021, p = .756 \) (reaction time).

In half the trials, a road image was then presented for 2000ms and participants were asked to verbally identify the hazard. The experimenter recorded whether the response was correct or incorrect using the keyboard. The experimenter was trained to code anything that resembled the pre-determined (see materials) hazard as accurate and anything else as inaccurate. For example, for pedestrians stepping out into the road, participants could say "those people stepping out" or "pedestrians there" and pointing. The computer recorded the onset of the spoken response and this was taken as response time to the hazard. There was no time limit to respond, but all participants responded within 3s. After 2000ms the screen went blank, consistent with Thompson and Crundall (2011). In the other half of the trials a further two letter searches (in the same orientation as the first letter string) were presented before the road picture was shown. A total of 108 trials were completed by each participant; 36 for each orientation, with 18 incorporating one letter search and 18 incorporating three letter searches. The purpose of including different numbers of letter strings before the road image was to increase the unpredictability of the onset of the road image. This unpredictability is crucial in the creating the carryover effect (Hills et al., 2016). There were an equal number of trials in each condition with 3 or 4 vowels and with each road type. Trials were presented in a random order and road images were selected at random. Eye movements were recorded for the entire duration of the trial.

Once all 108 trials had been completed, participants were given a short demographic questionnaire which explored their driving experience. Finally, they were thanked and debriefed.

**Results**

Data collected were the hazard identification accuracy and response times, in addition to the proportion of horizontal and vertical eye movements during viewing of the road images. We excluded any trials in which participants did not count the number of vowels correctly (6% of trials). We first analysed the saccadic direction followed by the behavioural data. Table 2 summarises the results from this study.
Eye Movements

In order to replicate the basic effect of Thompson and Crundall (2011), we first analysed the saccade direction data employing an analytical structure similar to Gilchrist and Harvey (2006). The direction of each saccade was measured in degrees and saccades were coded into one of four 90 degree bins (zero degrees represented a vertical upwards saccade and 180 degrees represented a vertical downwards saccade). The resulting bins represented upward (covering 316° to 45°), downward (covering 126° to 225°), leftward (covering 226° to 315°), and rightward (covering 46° to 125°) saccades. All eye movement analyses were conducted after the initial fixation on the road: we discounted the first fixation due to contamination from the trial structure. Analysis of the eye-movement data was conducted only on the trials in which the hazard was identified correctly. Figure 1 represents the mean number of saccades made in each direction. In the subsequent analysis, we collapsed across left and right movements to create a single horizontal eye movement measure and across up and down to create a single vertical eye movement measure. In this way, we could represent 1 as vertical eye movements and 2 as horizontal eye movements and estimate the amount of horizontal and vertical eye movements made per trial. Mean eye movement score was analysed in a 3 x 3 x 3 mixed measures ANOVA with the within-subjects factors of letter string orientation and hazardousness of the road, and the between-subjects factor of driving experience. Table 2 represents the data as percentages of horizontal eye movements for clarity.
Figure 1. Mean spread of saccades following the letter strings split by road hazardousness for Experience Drivers (top panel), Novice Drivers (middle panel), and non-drivers (bottom panel). Values represent the mean number of saccades in each direction.
Table 2.

Mean (and standard error) hazard detection RT (ms), hazard perception accuracy (%), hazard rating, and amount of horizontal eye movements (%) as a function of letter string orientation, hazardousness of road and driver experience.

<table>
<thead>
<tr>
<th>Orientation of letter string and Hazardousness of road</th>
<th>Non Driver</th>
<th>Novice Driver</th>
<th>Experienced Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazard detection RT (ms)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Hazard</td>
<td>1272 (50)</td>
<td>1173 (42)</td>
<td>944 (57)</td>
</tr>
<tr>
<td>Medium Hazard</td>
<td>1201 (46)</td>
<td>1107 (39)</td>
<td>933 (52)</td>
</tr>
<tr>
<td>High Hazard</td>
<td>1229 (44)</td>
<td>1090 (37)</td>
<td>951 (50)</td>
</tr>
<tr>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Hazard</td>
<td>1395 (50)</td>
<td>1274 (42)</td>
<td>1086 (57)</td>
</tr>
<tr>
<td>Medium Hazard</td>
<td>1387 (49)</td>
<td>1186 (41)</td>
<td>1040 (55)</td>
</tr>
<tr>
<td>High Hazard</td>
<td>1377 (56)</td>
<td>1204 (48)</td>
<td>1029 (64)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Hazard</td>
<td>1332 (43)</td>
<td>1160 (36)</td>
<td>1029 (49)</td>
</tr>
<tr>
<td>Medium Hazard</td>
<td>1297 (45)</td>
<td>1171 (38)</td>
<td>997 (51)</td>
</tr>
<tr>
<td>High Hazard</td>
<td>1275 (45)</td>
<td>1185 (39)</td>
<td>1002 (52)</td>
</tr>
<tr>
<td>Hazard detection accuracy (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>83.77 (2.53)</td>
<td>79.83 (2.30)</td>
<td>91.48 (2.89)</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>85.49 (2.40)</td>
<td>88.49 (2.04)</td>
<td>93.48 (2.63)</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>74.94 (2.97)</td>
<td>76.60 (2.52)</td>
<td>93.89 (2.74)</td>
</tr>
<tr>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>78.40 (2.61)</td>
<td>79.42 (2.22)</td>
<td>90.48 (3.24)</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>80.00 (2.72)</td>
<td>83.58 (2.50)</td>
<td>92.05 (3.10)</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>82.21 (2.95)</td>
<td>82.99 (2.50)</td>
<td>82.82 (3.36)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>80.30 (2.46)</td>
<td>82.59 (2.09)</td>
<td>92.91 (2.80)</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>86.14 (2.32)</td>
<td>88.65 (1.97)</td>
<td>91.62 (2.64)</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal eye movements (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>58.76</td>
<td>63.26</td>
<td>76.64</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>55.95</td>
<td>66.03</td>
<td>79.87</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>56.48</td>
<td>65.18</td>
<td>80.97</td>
</tr>
<tr>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>20.92</td>
<td>35.49</td>
<td>39.11</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>23.92</td>
<td>33.10</td>
<td>52.94</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>21.77</td>
<td>34.00</td>
<td>57.50</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Driver</td>
<td>43.34</td>
<td>59.24</td>
<td>74.32</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>42.88</td>
<td>59.68</td>
<td>83.00</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>43.10</td>
<td>61.55</td>
<td>83.69</td>
</tr>
</tbody>
</table>
This analysis revealed that the letter string orientation affected eye movements, $F(2, 158) = 95.56, MSE = 0.06, p < .001, \eta_p^2 = .55$. Specifically, there were more horizontal eye movements following the horizontal ($p < .001$) and random letter strings than the vertical letter strings ($p = .229$). There were more vertical eye movements following the vertical letter search than the random letter search ($p < .001$), replicating Thompson and Crundall (2011).

The main effect of driver experience was significant, $F(2, 79) = 25.77, MSE = 0.22, p < .001, \eta_p^2 = .40$, whereby experienced drivers displayed more horizontal eye movements than novice ($p < .001$) and non-drivers ($p < .001$). Novice drivers displayed more horizontal eye movements than non-drivers ($p < .001$). The main effect of hazardousness of the road image was also significant, $F(2, 158) = 4.00, MSE = 0.03, p = .020, \eta_p^2 = .05$. There were fewer horizontal eye movements made to low hazardous images relative to high hazardous images ($p = .016$).

There was an interaction between driver experience and hazardousness of the road, $F(4, 158) = 3.25, MSE = 0.03, p = .014, \eta_p^2 = .08$. Three one-way ANOVAs were conducted: one for each level of driving experience with the hazardousness of the road image used as the independent variable. This revealed no significant effect of hazardousness of the image on scanning behaviour of non-drivers, $F(2, 50) = 0.02, MSE = 0.01, p = .982, \eta_p^2 < .01$, and novice drivers, $F(2, 70) = 0.09, MSE = 0.01, p = .912, \eta_p^2 < .01$. Hazardousness did influence the scanning behaviour of experienced drivers, $F(2, 38) = 6.88, MSE = 0.02, p = .003, \eta_p^2 = .27$, with fewer horizontal eye movements toward the least hazardous images relative to the medium ($p = .020$) and the most hazardous images ($p = .003$). This replicates Hughes and Cole (1986). These findings indicate that drivers of greater experience are more likely to modify their strategy based on the level of hazardousness.

**Hazard Perception Response Time**

To assess whether the letter string orientation affected hazard perception, as found by Thompson and Crundall (2011), we ran a parallel analysis on the response time to the hazard identification. This revealed a main effect of letter string orientation, $F(2, 158) = 34.21, MSE = 32654, p < .001, \eta_p^2 = .30$, in which road hazards were identified faster following the horizontal letter strings than the random (mean difference = 61ms, $p < .001$) and the vertical letter strings (mean difference = 120ms, $p < .001$). Responses were slower following the vertical letter strings than the random letter string (mean difference = 59ms, $p < .001$). These results replicate Thompson and Crundall (2011). The main effect of driver experience was significant, $F(2, 79) = 12.35, MSE = 385589, p < .001, \eta_p^2 = .24$, replicating McKenna and Crick (1999). Experienced drivers were significantly faster at responding to hazards than
non-drivers (mean difference = 306ms, \( p < .001 \)) and novice drivers (mean difference = 171ms, \( p = .012 \)). Novice drivers were faster than non-drivers at responding to the hazard (mean difference = 135ms, \( p = .040 \)). There was also a main effect of hazardousness of the road, \( F(2, 158) = 6.15, MSE = 17450, p = .003, \eta^2_p = .07 \), in which responses were faster to the most hazardous compared to the least hazardous roads (mean difference = 36ms, \( p = .018 \)) and faster to the medium than the least hazardous roads (mean difference = 38ms, \( p = .008 \)). There was no difference in response times to the high and medium hazardous roads (mean difference = 2ms, \( p = 1.00 \)). No interactions were significant (largest \( F = 1.32 \), smallest \( p = .235 \)).

**Hazard Perception Accuracy**

Extending the effects of letter string reading onto hazard identification accuracy, and replicating Hills et al. (2016), we ran a parallel analysis on the hazard identification accuracy data. This revealed a main effect of letter string orientation, \( F(2, 158) = 15.32, MSE = 0.01, p < .001, \eta^2_p = .16 \). Hazard identification was more accurate following the horizontal letter string than the random (mean difference = 2%, \( p = .016 \)) and vertical letter string (mean difference = 5%, \( p < .001 \)). Vertical letter strings led to lower hazard perception performance than the random letter strings (mean difference = 3%, \( p = .013 \)). The main effect of driver experience, \( F(2, 79) = 6.18, MSE = 0.08, p = .003, \eta^2_p = .14 \), revealed that experienced drivers were more accurate at hazard identification than novice (mean difference = 8%, \( p = .016 \)) and non-drivers (mean difference = 10%, \( p = .004 \)). There was no significant difference in hazard identification accuracy between the novice drivers and the non-drivers (mean difference = 2%, \( p = 1.00 \)) indicating that this hazard identification task may have lacked sufficient sensitivity (due to the requirement to verbalise the hazard). The main effect of hazardousness of the road was also significant, \( F(2, 158) = 16.06, MSE = 0.01, p < .001, \eta^2_p = .17 \). Hazard identification accuracy was higher for the high hazard images than the medium (mean difference = 3%, \( p = .003 \)) and the low hazard images (mean difference = 6%, \( p < .001 \)). Hazard identification accuracy was also higher for the medium hazard images than the low hazard images (mean difference = 3%, \( p = .020 \)).

In this analysis, we also found a three-way interaction, \( F(8, 316) = 2.44, MSE = 0.01, p = .014, \eta^2_p = .06 \). In order to explore this interaction, we ran three separate 3 x 3 within-subjects ANOVAs with the letter string orientation and hazardousness of the road as factors. We looked at the effect of the letter string and the interaction between the letter string orientation and the road hazardousness at each of the levels of driver experience. The results showed that the effect of the letter string was smaller for the experienced drivers, \( F(2, 38) =$
3.76, $MSE = 0.01$, $p = .032$, $\eta_p^2 = .17$, than the novice drivers, $F(2, 70) = 11.60$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .25$, and non-drivers, $F(2, 50) = 6.88$, $MSE = 0.01$, $p = .002$, $\eta_p^2 = .22$. The interaction between hazardousness of road and letter string orientation was significant only for the experienced drivers, $F(4, 76) = 3.28$, $MSE = 0.01$, $p = .032$, $\eta_p^2 = .15$. This interaction was revealed through a non-significant effect of letter string orientation for the high hazard images on accuracy, $F(2, 38) = 0.42$, $MSE = 0.01$, $p = .663$, $\eta_p^2 = .02$, whereas it was for the medium, $F(2, 38) = 6.45$, $MSE < 0.01$, $p = .004$, $\eta_p^2 = .25$, and low hazard images, $F(2, 38) = 4.53$, $MSE = 0.01$, $p = .017$, $\eta_p^2 = .19$. These results highlight that the carryover effect exists for all participant groups, but the effect on hazard perception is smaller for the experienced drivers and importantly, on the hazardous roads, the experienced drivers are less affected by vertical carryover on their hazard performance.

We also correlated hazard perception accuracy and response time with number of horizontal saccades to further demonstrate the relationship between scanning behaviour and hazard perception. This revealed a significant positive correlation between amount of horizontal scanning and hazard identification accuracy, $r(80) = .33$, $p = .003$, and a significant negative correlation between the amount of horizontal scanning and hazard identification response time, $r(80) = -.55$, $p < .001$.

Secondary Analysis: Task Difficulty
The results found thus far are consistent with previous research. We have replicated the carryover effect found by Thompson and Crundall (2011) and Hills et al. (2016) showing that eye movements to road scenes are influenced by the layout of stimuli in a preceding task. In addition, we have further evidence showing that the carryover of vertical eye movements appears to be detrimental to hazard perception with increased response times and reduced accuracy when identifying a hazard. We have found that this carryover of eye movements occurs for drivers of different levels of experience and non-drivers alike, but that experienced drivers are less affected by the carryover in terms of hazard identification accuracy and response time.

In the introduction, we posited either that experienced drivers should have a narrower attentional spotlight (Lavie et al., 2004) and be able to ignore distractors more easily than novice drivers, or experienced drivers may find distractions harder to ignore than novice drivers when the task itself is demanding (McKenna and Farrand (1999) because experienced drivers behaviours are more malleable. Thus far, our results have indicated that the carryover effect does not influence the responses to hazardous scenes more than non-hazardous ones.
However, there is no correlation between hazardousness and ease of detection of the hazard found in previous work (McGowan & Banbury, 2004), nor in our current stimuli set, $r(80) = .04, p = .702$. Therefore, to test whether increased task demands, operationalised as the difficulty with which the hazard is to detect, affects the magnitude of eye movement carryover, we reran the analyses but using stimulus difficulty as an independent variable, replacing stimulus hazardousness. From the current dataset, we established in which stimuli the hazards were harder to detect. Note that the identification of hazard identification difficulty was established from the average of all participants' performance of images in random letter-string orientation (as this best represented a control condition) rather than across all stimuli as this could have been affected by the letter string orientation (see methods). Any stimulus with hazard identification accuracy of below 75% was considered a difficult stimulus. Any stimulus with hazard identification accuracy of above 96% was considered an easy stimulus, and those in between were considered a moderate level of difficulty.

We ran a parallel set of analyses to those described above. For this summary, we only report the results involving the factor of difficulty as all other effects were as reported above. Firstly, for saccadic direction, we found that difficulty affected eye movement direction, $F(2, 146) = 8.55, MSE = 0.04, p < .001, \eta_p^2 = .11$, in which there were more horizontal eye movements made when the hazard identification was easy than when it was moderately ($p = .080$) or very difficult ($p = .001$). There was a non-significant trend for more horizontal eye movements when the hazard identification was moderately difficult relative to very difficult ($p = .096$). There was also a three-way interaction between driver experience, letter string orientation, and difficulty of stimulus, $F(8, 292) = 2.88, MSE = 0.03, p = .004, \eta_p^2 = .07$. We conducted three 3 x 3 within-subjects ANOVAs to decompose this interaction, one for each group of participants. These revealed a significant interaction between letter string orientation and difficulty for the non-drivers, $F(4, 100) = 6.70, MSE = 0.03, p < .001, \eta_p^2 = .21$, but not the novice, $F(4, 100) = 1.83, MSE = 0.05, p = .127, \eta_p^2 = .06$, nor experienced drivers, $F(4, 100) = 0.86, MSE = 0.02, p = .492, \eta_p^2 = .05$. The main effect of stimulus difficulty was significant for the experienced drivers, $F(2, 34) = 4.50, MSE = 0.02, p = .018, \eta_p^2 = .21$, and the non-drivers, $F(2, 50) = 7.83, MSE = 0.04, p = .001, \eta_p^2 = .24$, but not the novice drivers, $F(2, 62) = 1.02, MSE = 0.05, p = .367, \eta_p^2 = .03$. Both main effects were as described above: more horizontal fixations for the easy stimuli relative to the moderately and highly difficult stimuli. These results imply that the experienced drivers were more affected by task difficulty than the novice drivers replicating McKenna and Farrand (1999). Furthermore, non-drivers
were affected by the hazards potentially due to a lack of experience searching hazards (having not performed hazard perception on a regular basis).

Difficulty affected hazard identification response time, $F(2, 146) = 183.29$, $MSE = 22853$, $p < .001$, $\eta_p^2 = .72$: Easier hazards were identified faster than moderately (mean difference = 70ms, $p < .001$) and very (mean difference = 230ms, $p < .001$) difficult hazards. Very difficult hazards were identified slower than moderately difficult hazards (mean difference = 160ms, $p < .001$). There was no interaction between task difficulty and driver experience, $F(4, 146) = 0.75$, $MSE = 22853$, $p = .517$, $\eta_p^2 = .02$.

In terms of hazard identification accuracy, the main effect of difficulty was obviously significant as we had determined the level of difficulty based on the present dataset, $F(2, 148) = 133.16$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .64$. Of interest was the significant three-way interaction, $F(8, 296) = 2.98$, $MSE = 0.01$, $p = .007$, $\eta_p^2 = .08$. Three 3 x 3 within-subjects ANOVAs were run - one for each participant group. These revealed a significant interaction between letter string orientation and difficulty for the non-drivers, $F(4, 124) = 4.45$, $MSE = 0.02$, $p = .009$, $\eta_p^2 = .13$, and the novice drivers, $F(4, 100) = 2.61$, $MSE = 0.01$, $p = .049$, $\eta_p^2 = .09$. However, it was not significant for the experienced drivers, $F(4, 72) = 0.09$, $MSE = 0.01$, $p = .967$, $\eta_p^2 = .01$. These interactions were revealed through bigger differences in hazard identification accuracy for the easy and difficult images following the horizontal and vertical letter strings than the random letter strings. This indicates that the impact of the letter strings compounded the identification of hazards for the non- and novice drivers but not the experienced drivers, which is somewhat contrary to the notion that experienced drivers would be more distracted by a dual task. These results are shown in Table 3.
Table 3.

Mean (and standard error) hazard perception accuracy (%) and amount of horizontal eye movements (%) as a function of letter string orientation, hazard identification difficulty and driver experience.

<table>
<thead>
<tr>
<th>Orientation of letter string and Hazard Identification Difficulty</th>
<th>Non Driver</th>
<th>Novice Driver</th>
<th>Experienced Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazard detection accuracy (%)</td>
<td>Easy Image</td>
<td>Horizontal Moderate Image</td>
<td>Hard Image</td>
</tr>
<tr>
<td>Non Driver</td>
<td>98.59 (1.28)</td>
<td>88.23 (1.64)</td>
<td>61.43 (3.59)</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>94.72 (1.42)</td>
<td>89.72 (1.82)</td>
<td>79.32 (3.98)</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>97.85 (1.66)</td>
<td>95.47 (2.13)</td>
<td>85.88 (4.66)</td>
</tr>
<tr>
<td>Horizontal eye movements (%)</td>
<td>Non Driver</td>
<td>52.26</td>
<td>46.80</td>
</tr>
<tr>
<td>Novice Driver</td>
<td>64.76</td>
<td>59.80</td>
<td>56.23</td>
</tr>
<tr>
<td>Experienced Driver</td>
<td>72.72</td>
<td>67.89</td>
<td>68.75</td>
</tr>
</tbody>
</table>
Discussion

This study replicates and extends the findings of Thompson and Crundall (2011), whereby the scanning strategy used in one task transfers to a second task (despite the two tasks being unrelated). They found increased vertical scanning in driving stimuli following a vertical letter search (compared to a random letter search) which has been replicated here. Numerically, horizontal eye movements increased following a horizontal letter search relative to following a random letter search however this effect was not significant (as in Thompson & Crundall, 2011). Horizontal search in driving scenes may be less influenced by a preceding task because attention would be directed in this manner on the basis of top-down influences (Crundall & Underwood, 1998).

Consistent with Mourant and Rockwell (1972; see also, Konstantopoulos et al., 2010; Underwood et al, 2002), we also found that experienced drivers searched horizontally more and vertically less than novice and non-drivers. Experienced drivers also modified their scanning behaviour based on the hazardousness of the road, displaying more horizontal scanning for the more hazardous roads (consistent with Chapman et al., 2002; Crundall & Underwood, 1998; Falkmer & Gregersen, 2001). However, eye movement carryover was effectively the same for all participants.

The most important finding from this study for real-world driving scenarios is that a preceding letter string influences how quickly and accurately participants detect hazards as a consequence of the carryover of scanning behaviour. Following horizontal searching, participants identified hazards faster and more accurately than following vertical or random searching. The horizontal condition may have resulted faster responses due to the fact that eye movements during driving are more focused in the horizontal plane (Crundall & Underwood, 1998); therefore ensuring a more appropriate search for detecting hazards. Vertical searching led to poorer hazard identification performance, supporting this argument. Experienced drivers were quicker and more accurate at identifying hazards, consistent with Crundall et al. (2003). We did find differential effects of the letter string orientation on drivers of different experience for hazard identification accuracy (but not response time): experienced drivers were less affected by carryover in hazard identification accuracy. This is inconsistent with the eye movement analysis described above. This indicates that experienced drivers are able to compensate for the potential distracting effect of the carried over eye movement pattern.
To assess how task difficulty modulated the carryover effect, we measured the carryover effect in stimuli in which the hazard was easy, moderately, or very difficult to identify. We found that the amount of horizontal scanning decreased as the task demands increased, consistent with the notion that horizontal scanning is based on mental effort and when the task demands are too great, they are reduced (Recarde & Nunes, 2003). The scanning pattern of experienced drivers was more affected by the increased task demands than the novice drivers. A further advancement to the Thompson and Crundall (2011) work is that here we have demonstrated that the detrimental effect of the vertical letter search on hazard identification accuracy was greater when viewing more demanding road scenes. Hazard perception is more affected by the preceding task when the primary driving task is more difficult. This result indicates that precisely when performance is vital for the driving, it is more at risk of being affected by extraneous tasks. Indeed, this is consistent with the notion of a limited resource attentional system that uses short-cuts when faced with more challenging demands (Shiv, & Fedorikin, 1999). In this case, the "short-cut" is unavoidably not readjusting previous attentional settings between tasks. This was the case for the novice and the non-drivers, but not the experienced drivers, who were able to counteract the detrimental effect of carryover, contrary to McKenna and Farrand (1999). Potentially because the experienced drivers can adapt their search according to the road anyway they adopt a more flexible attentional style and so they do re-adjust, yet novices have a more static allocation of attention so are not flexible in responding to the different tasks and environments.

The overall summary of these results is that the way in which participants allocated their attention and search to the driving scenes was influenced by the layout of stimuli in a preceding, secondary task, and when the task was more difficult, experienced drivers' eye movements were more affected by the carryover. However, in the behavioural performance, the effect of carryover was greater for novice drivers than experienced drivers. This result demonstrates that experience allows for compensatory strategies to be employed such that the effect of the secondary task is not as damaging to the critical aspects of hazard perception. We can use Groeger's (2000) theorising about why experienced drivers are more accurate at hazard perception more generally than novice drivers to explain these results. Due to their knowledge, experienced drivers have encountered more hazardous traffic situations and have stored these in memory. This means that any new encounter can be processed quickly and efficiently by an automatic cued recall of the previous incident. Since novice drivers have less experience of hazardous driving situations by virtue of not having driven so much, they
cannot rely on memory recall in order to process a driving scene. Instead, they have to engage in more effortful processing than their experienced counterparts. While the carryover effect may influence the eye movements of experienced drivers, they are still able to recall a similar hazardous situation and therefore react appropriately even with less efficient visual input. This fits with the notion that, due to a lack of experience, novice drivers are less willing to classify images as hazardous than experienced drivers (Wallis & Horswill, 2007). A further benefit afforded to experienced drivers over novices is that they can utilise covert and overt attentional strategies to hazard perception (Underwood 1998, Underwood et al. 2002, 2005): only overt attentional strategies may be affected by the carryover of eye movements from a preceding task.

In the preceding paragraph, we have highlighted another aspect of novice driving behaviour - they are more affected by a secondary task than experienced drivers. The resulting behavioural effect on hazard identification was greater following the same amount of carryover as in experienced drivers. This means that any in-car display (such as a SatNav or mobile phone) that utilises vertical search will potentially be causing more danger to novice drivers than experienced drivers. Given that novice drivers are more likely than experienced drivers to be involved in collisions than experienced drivers (Cooper et al., 1995) caused by distraction (Klauer et al., 2006; Stutts et al., 2001), this result has even more real-world implication. Hazard perception (McGowan & Banbury, 2004; McKenna & Crick, 1994; McKenna, Horswill, & Alexander, 2006) and eye movement training has been found to be effective in training participants to make more horizontal eye movements (Chapman et al., 2002) which is beneficial to hazard perception. Potentially, such training could be developed to limit the effects of detrimental vertical carryover.

Thompson and Crundall (2011) indicated that the mechanism for this carryover effect may be accounted for using theory of attentional weighting developed by Bundesen (1990). Attentional weights are allocated to particular locations by top-down (e.g., Gilchrist & Harvey, 2006; Loftus & Mackworth, 1978), bottom-up (e.g., Itti & Koch, 2000, Torralba et al., 2006), and preceding task cues. Attentional weights from one task persist to the subsequent task where they interact with weights allocated on the basis of saliency and past experience. Therefore, in the current scenario, weights are allocated to the positions of the letters in the letter string task and these weightings persist, at least for a short time (Thompson & Crundall's, 2011, data indicate this effect lasts a maximum of 2 seconds) in the subsequent task. The explanation requires a separation between abstract spatial locations and scenic locations: attentional weights are given to left or right regions of the next image,
whatever they may be. An alternative explanation is that the effect is simply an oculomotor reflex persisting from a highly skilled task (reading) for a few seconds before being inhibited, though this latter explanation seems to have been ruled out by Hills et al. (2016).

The present results, coupled with those of Thompson and Crundall (2011; Hills et al., 2016; and Thompson et al., 2015) add to Ball and Owsley's (1991) predictive model of unsafe drivers. In their model, they indicate that processing speed, divided attention, and selective attention combine to indicate a useful field of view that drivers adopt to complete the task of driving (Owsley, Ball, Sloane, Roenker, & Bruni, 1991; Owslet, Stalvey, Wells, Sloane, & McGwin, 2001). A secondary task will impair drivers useful field of view (see e.g., Ball, Beard, Roenker, Miller, & Griggs, 1998). Hills et al. (2016) have demonstrated that the carryover of eye movements is unrelated to processing speed and divided attention. Furthermore, its relation to selective attention is fairly minimal suggesting that the metric used to calculate the useful field of view would be enhanced by incorporating elements of this carryover effect. This will further increase its predictability of who will be involved in road accidents.

One concern with the results of this study is that participants were expected to detect and identify hazards in every trial. This task is unlike real-world driving. The present study should therefore be replicated when hazards are not present in every trial. Furthermore, the effect of carryover of eye movements should be assessed in different driving tasks such as lane control as these may also be affected by eye movement patterns.

The results of this study show a highly important effect: carryover of eye movements from an irrelevant letter search interferes with hazard identification and the allocation of attention. Following searching horizontally, subsequent hazard perception performance is faster and more accurate than that following a random or a vertical search. Anything that interferes with visual search of a driver is therefore a detriment to hazard perception. The effect of carryover is more detrimental to novice drivers than experienced drivers and potentially may contribute to their involvement in more accidents. The carryover of eye movements between two unrelated tasks is therefore an important effect which warrants further investigation.
References


Transport and Road Research Laboratory (1979). A hazard perception test for drivers, Leaflet. Transport and Road Research Laboratory, Crowthorne, UK.

