ABSTRACT

The SAFEye™ video eye tracking methodology was used to objectively evaluate sign conspicuity and driver reaction times to a lane closure on a motorway. A single blind experimental design with 24 participants comprising four different scenarios depicting a static lane closure layout without advanced attenuator trucks were combined with memory recall questions and evaluated under laboratory conditions. Drivers first observed a sign about 110 m away, but only semantically processed its meaning at about half (64m) of the initial distance from first observing a sign. This allows only a further 30m (1s), to process the meaning before it becomes illegible travelling at 100km/h. The average total duration spent processing a dynamic object using pursuit eye fixations is around 400ms. Although 100% of drivers initially gazed at the first TW-7 sign, fewer than 80% actually fixated it longer than 75ms. Correct identification of the lane closure sign (TW-7), was 67% when asked after 32s, and for the speed reduction (RG-4) sign, only 63% recalled it accurately after 11s. However, this accuracy deteriorated to 42% of the drivers correctly recalling what action, such as changing lanes, should be taken, and only 33% could remember after 11s what speed they should travel at.
BACKGROUND

Human perception of visual information is a critical factor in determining the safe use of both the vehicle and the interacting environment. A driver's perception reaction time under different driving conditions comprises perception of the stimuli, in this case a visual stimuli, the cognitive process, and the resulting action(s). In addition, a driver's anticipation of the event, such as searching, is influenced by other attention grabbing events such as hazards (Underwood et al., 2003), delaying detection of target signs by up to 1.5s after a hazard being observed. Edquist (2008) found that changes to road signs were harder to detect than changes to other vehicles, with roadside clutter significantly affecting the speed of detection for changing signs, which is consistent with the visual search literature. The human visual field can be broken down into the foveal vision (the central 2˚); the parafoveal (extends 5˚ either side of the central field) and the peripheral field (about 95˚ either side from the central field) (Rayner and Pollatsek, 1992). When we view a visual scene, assimilation of the contents is built up by a succession of fixations with the eye held virtually motionless at a fixed position for the duration of typically a few hundred milliseconds, allowing localized detail of the scene to be gathered through the retina preferentially in the region of the fovea zone (Rayner, 2009, Rayner, 1998, Rayner and Pollatsek, 1992). Visual acuity drops off very rapidly with increasing retinal eccentricity, being reduced by 50% at a distance of 5˚ from the fovea and by 90% at 40˚ (Hochberg, 1978) cited in Pollatsek and Rayner (1992).

A recent review of how humans perceive and retain scene information reveals conflicting theories with regards to how much visual memory, if any, do we possess. Transsaccadic visual theory states that humans perceive a scene by combining detailed information from the scanning of the scene by a succession of fixations. (Pollatsek and Rayner, 1992, Henderson and Hollingworth, 2003) However, more recent theories appear to support the notion of a visual short term memory (VSTM) to help integrate scene information. It also is suggested that we have a perceptual awareness of around 4˚ when viewing a scene. (Rayner, 2009) Whichever theory is more explanatory, what is well known is that humans are mostly unaware of a large object, or even colour changes in scenes during saccades when the object is not the next target fixation. However, if the object is a target to be fixated preceding the next saccade, then the change is likely to be noticed. (Henderson and Hollingworth, 2003) During a saccade, there is a saccadic suppression, during which the visual field is essentially a blur, or smear across the retina, and the brain does not record any information during this time. (Rayner and Pollatsek, 1992, Irwin and Zelinsky, 2002). Research by Galpin (2009) discusses the phenomena of change blindness in the driving scene, where a person looks, but does not see. Change blindness manifests itself in driving when a change to a relevant object is missed through temporary occlusion by an obstruction, such as a passing vehicle, or during a foreground eye pursuit movement that tracks over the top of another background target object, also during saccades between fixations when scanning the scene for other important targets, even during glances to mirrors, or to the dashboard.

Previous research found that due to the dynamic nature of driving that the frequency of saccades (the eye movement between fixations) increase, thus any changes during saccades are very pertinent. However, in the case of dynamic driving scenes, signs and other features move from the centre field to the edges, expanding in size. This requires the eye to undertake a pursuit movement holding the object steady in the foveal view, but the eye fixation moves relative to the boundary of the field of view. The traditional definition of a static fixation is, therefore, not applicable, but should be measured as a pursuit fixation by combining the relatively short fixations within a defined area of object expansion. The SAFEye methodology uses this eye pursuit calculation to estimate how long a moving object is fixated using the minimum duration of 50ms as a differentiator between a saccade and a pursuit fixation. The average duration of fixation and a saccade for different tasks are presented in Table 1 (Rayner, 1998, Underwood et al., 2002).
These mean values are taken from different sources (Rayner, 1998), and despite that various methods of presentation and measurement were used to gather the data it does provide a general appreciation of how fixation durations vary between distinctly different visual tasks. The mean duration of affixation increases with task and scene complexity, to the point where a dynamic scene, as experienced by drivers, tends to require twice as long as simple reading tasks. Irwin & Zelinsky (2002) investigated the relationship between eye movements, scene perceptions and memory. They found that the accuracy is a function of whether the object was foveated or not, with foveated objects always reporting much higher recall accuracy, 92% for foveated objects versus 53% of non-foveated objects. They also found that our accuracy is significantly higher only for the last three objects inspected; with much poorer recall of objects fixated further back in time. Although scene information appears to be improved with more fixations, our ability to remember specifics in the scene is limited to about the most recent five items. This indicates our working memory is considerably limited when it comes to scene inspection. A study by Underwood et al. (2003) confirmed that hazardous and moving objects attract more fixations and attention than centrally located static objects. Hazardous objects impaired detections for up to 1.5s after onset of the hazard, during which viewers fixate longer, and more often on threatening objects. (Underwood et al., 2002)

This outlines the problems and misconceptions associated with human vision, in that despite our cinematic-like pseudo awareness of the scene, humans don’t retain much detail of the entire scene beyond our sensory memory (1-2s). This is a major problem when drivers are confronted with competing targets, insufficient time to process them fully, and then trying to remember the action that may be required in the near, but not immediate future. The corollary of providing too much information in the pursuit of safety is that the driver overload increases the risk of missing other vital information.

The purpose of this experiment was to evaluate the conspicuity of the signs by drivers in terms of cognitive awareness when approaching static warning, temporary and regulatory signs showing a left lane closure. This static layout is proposed as an alternative to the use of two advance warning attenuator trucks (Figure 1) with flashing lights and keep-right sign. Figure 2 shows the sequence of the five signs that were visible to the drivers during the experiment. The TW7 sign was varied to reflect 2 lanes (SC3&4) and 3 lanes (SC7&8). (See Table 1)
Figure 2: Signage visible to drivers

METHODOLOGY

Experimental Design: The scenarios were video recorded using a dashboard mounted video camera recording an approximate similar view as seen from the driver’s seated position, representing a forward view between the vehicle’s A-pillar and cabin rear view mirror. This arc of view spans about 45˚. The video recording starts well upstream from experimental segment that has been prepared in accordance with the Code of Practise for Temporary Traffic Management (COPTTM) layouts for a lane closure on a motorway. Signs were only placed on the left as a safe method for deployment of signs on the right side in the median could not be developed at the time.

<table>
<thead>
<tr>
<th>Recording</th>
<th>Scenario Reference Code</th>
<th>Left sign placement</th>
<th>Right sign placement</th>
<th>Lane change needed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-lane</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-lane 1</td>
<td>SC3</td>
<td>yes</td>
<td>no</td>
<td>Yes</td>
</tr>
<tr>
<td>Right-lane 2</td>
<td>SC4</td>
<td>yes</td>
<td>no</td>
<td>No</td>
</tr>
<tr>
<td>3-lane</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-lane 1</td>
<td>SC7</td>
<td>yes</td>
<td>no</td>
<td>Yes</td>
</tr>
<tr>
<td>Right-lane 3</td>
<td>SC8</td>
<td>yes</td>
<td>no</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 2: Experimental design

The scenarios were as set out in Table 1 using a repeated measures design with all subjects tested on all four scenarios. For both the two-lane and three-lane setup, recording took place in the far left lane and the far right lane to test the difference in observation and perceptual awareness between drivers that need to merge actively from the left lane, and those drivers that need to be made aware that drivers will be forced to merge from the left.

The experiment tested driver perceptual awareness through an assessment of the gaze intensity, fixation duration and perceptual awareness of the TW- signage. In order to eliminate response, fatigue or practice bias, random selection of one of two different sequences of video clips with associated questions and dummy videos of urban roads were presented to the subjects. Sequence 1 showed; dummy1, dummy2, dummy3, SC3, SC4, SC7 and SC8. Sequence 2 showed; SC3, dummy1, SC7, dummy 2, SC4, dummy3, SC8 and dummy4.

Apparatus: The video footage was taken with a dash mounted iPhone 4S high definition video camera at 29 fps and 1920 x 1080 resolution. The view was matched to that approximately from the driver’s viewpoint between the A pillars with the top boundary adjusted to match that limited by the cabin roofline. The videos were post-processed, trimmed and spliced together with other footage into the EyeMetrix Recording software. The video, subject instructions and questionnaire were developed in the script file for testing. The testing took place under controlled conditions using a Mirametrix S2 eye tracker with a 0.5 to1 degree accuracy sampling at 60 Hz, to record eye movements, and the EyeMetrix OEM software and SPSS was used to analyse the data. The dynamic high definition video was displayed at 29 fps on a 23” widescreen at a 1920 x 1080 resolution placed approximately 600mm from the driver with the eye tracker placed at the bottom of
the screen, and a secondary screen was used to monitor the experiment. The field of view spanned 45° (horz.) and 25° (vert.) and the display screen of 500 mm x 280mm corresponds well with the driver’s field of view inside the car cabin between the cabin mirror and the A-pillar (500mm x 300mm at a 600mm distance).

**Participants:** A sample of 30 drivers, most holding full licenses with the exception of one on a learners, were recruited that represented a wide age range from 18 to 56 yrs. old (M age=36.33, SD age=10.04) and driving experience (M exp=18.41, SD exp=7.41). A final 24 were used in the analysis yielding a sufficient sample to apply the central limit theorem and is of similar scale to other relevant research studies (Charlton, 2006, Konstantopoulos et al., 2010, Borowsky et al., 2008). All subjects were given the same basic instructions of what to expect during the experiment. The experiment complied with the conditions obtained from the University of Auckland ethics committee.

**Procedure:** All subjects were briefed on what they should expect during the test, but were not told the purpose of the test other than to behave as if they were the driver of the video sequence and to answer multiple choice questions at the end of some of the sequences. The experimental setup included a games controller steering wheel and mouse to immerse the subjects into behaving more like a driver, as opposed to a passenger just viewing the video. The drivers were instructed to mimic the car movement and click the mouse when they felt a lane change was required. Upon completion of all testing, the data was imported into the EyeMetrix Analyser software, scanned and refined for further analysis. The guideline1 for sign visibility in a level 3 work zone is 120m with two TW signs spaced at 200m intervals, and the RG at 100m intervals. The data were analysed at the 200, 150, 120, 90, 60 and 30m positions from the sign, the 30m intervals correspond with approximately 1s of travel between each analysis slice. Closer than 30m, the angle between the sign and central line will increase to more than 10° and requires excessive rotation of the eye and/or head to view, and travelling at 100 km/h, it also becomes illegible at angles greater than 45° from the forward view.

**RESULTS**

**First time to fixation:** The time to first gaze (notice an object with incomplete processing) and fixation (complete processing of object meaning) is measured as the elapsed time in seconds from the beginning of the task until the first observation is recorded in the target region. Fixations are computed according to gazes that fell within 40 pixels and > 75ms of each other. The first time to fixation (M=3.108s, or 64m @ 100km/h from the sign) on to a TW-7 sign is greater than that for gazes (M=1.414s, or 111m @ 100 km/h from the sign). There is a significant difference between the time to first gaze and time to first fixate on the TW-7 sign, F (1,136) = 45.44, p<0.001. This means that there is a significant difference in time when drivers first observe that an object of interest should be closer inspected, and when they actually use a pursuit fixation to process the information at a later stage. Essentially, drivers delay the inspection until they are about halfway between when they first notice something of interest and the actual target, and when they actually pay attention to its meaning. There are no significant differences of fixations or gaze times between scenarios, meaning that the number of lanes, or lane position of the vehicle, did not influence the respective gaze or fixation times.

**Fixation Duration:** The majority of subjects (88-100%) gazed on the first TW-7 sign with little difference between scenarios. (Table 3) However, when drivers are in the left hand lane and have to change lanes due to the closure (SC3 and 7), 54% and 79% then respectively fixated the first TW-7 sign. Whereas, when the driver is in the right hand lane and the sign is irrelevant to the driver, the proportion drops to only 38% - 79% (SC4 & 8). This could be indicative of more experienced drivers shedding their attention on a sign that is deemed irrelevant, and the novices

---

still processing its meaning regardless of whether they are required to act upon it. Should this be the case then it appears that some semantic meaning is extracted during the much shorter gaze duration than currently thought (<75ms) to help decide whether or not to fixate it for longer when they are about 2s from the sign.

<table>
<thead>
<tr>
<th>Lane and position</th>
<th>Scenario</th>
<th>Time to first fixation (% subjects)</th>
<th>Fixation duration (ms)</th>
<th>Total fixation duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane 1 of 2</td>
<td>SC3</td>
<td>100%</td>
<td>54%</td>
<td>155</td>
</tr>
<tr>
<td>Lane 2 of 2</td>
<td>SC4</td>
<td>92%</td>
<td>38%</td>
<td>148</td>
</tr>
<tr>
<td>Lane 1 of 3</td>
<td>SC7</td>
<td>96%</td>
<td>79%</td>
<td>149</td>
</tr>
<tr>
<td>Lane 3 of 3</td>
<td>SC8</td>
<td>88%</td>
<td>46%</td>
<td>174</td>
</tr>
</tbody>
</table>

Table 3: % of subjects that fixated TW-7 sign

The range of single fixation durations on the TW-7 sign varies between 148 – 174ms. A one-way ANOVA with post-hoc tests show there are no significant differences between scenarios. The fixation duration between different areas of interest such as car following and opposing traffic is also non-significant. The mean total fixation duration on the TW7 sign over a distance of 200m for all scenarios is 400ms. There are no significant differences between scenarios for total fixation duration for the TW-7 sign, meaning that lane position or number of lanes did not influence driver fixation times.

Sign recognition: Drivers had to answer questions at the end of the video sequence about what signs they recognised and in a later sequence what actions they needed to take (Table 4).

<table>
<thead>
<tr>
<th>Memorability</th>
<th>Correct response</th>
<th>Time lag between sign visible and response required (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognised lane closed sign (TW-7)</td>
<td>67%</td>
<td>32</td>
</tr>
<tr>
<td>Recalled action to change lanes (TW-7)</td>
<td>42%</td>
<td>45</td>
</tr>
<tr>
<td>Recognised speed limit change (RG-4)</td>
<td>63%</td>
<td>11</td>
</tr>
<tr>
<td>Recalled action to reduce speed (RG-4)</td>
<td>33%</td>
<td>&gt; 11</td>
</tr>
</tbody>
</table>

Table 4: Action, memory and recognition of signs

As expected there is a reduction in accuracy over time when ask to recognise a sign passed earlier when presented with a list of possible signs, but there was a marked reduction in accuracy of what they can remember what action they were to take. Drivers tended to remember that there was a sign, but could not remember what action was required if an action were not applied immediately. This could be a result of incomplete processing, or load shedding to attend to more recent information being presented, such as lane keeping, guidance signage reading or attending to surrounding traffic. This reinforces the notion that if attenuator trucks are placed too far in advance from the actual site or far apart, as is often the case based on casual observation on the Auckland network, and no immediate action to the keep right sign is needed based on the drivers’ forward visibility to reinforce it, then the message is lost and their usefulness tending to irrelevant. This often manifests itself when drivers pass the first truck and after a while move back into the lane only to be confronted with the second truck when it becomes visible.

CONCLUSIONS

Static lane closure signage placed only on the left side of a 2 and 3 lane dual carriageway appear to be effective in alerting drivers on the motorway to a lane closure ahead, without the need of additional attenuator trucks that are placed upstream of the static signage.

Consistent with COPTTTM guidelines, drivers first observe a static sign about 110m away, and only begin to process its message around 65m from the sign. The majority of drivers actually looked at the first TW-7 sign from the left hand lane, but only between 55 and 78% actively processed its
information. Data for the second TW-7 sign still requires analysis to understand the total sign conspicuity.

Individual mean fixation durations for drivers inspecting both essential and non-essential features were around 150ms regardless of lane position or presence of road signage. Total mean pursuit fixation durations on the TW-7 sign is around 400ms which is consistent with findings by Underwood et al. (2002).

Of the 24 subjects tested in 4 scenarios on a 2-lane and 3-lane section where 67% correctly identified the TW-7 sign to change lanes when presented with pictures of the sign from the video, but only 42% could remember without sign priming what action they should have taken 45s after first observing the sign. Recall of a speed change signs was 63% correctly recognising the sign, but only 33% could remember that they had to reduce speed.

The findings are consistent with past studies and underpin the need for keeping the lag between signage and action as short as possible to reduce the opportunity for memory decay and incorrect actions.

REFERENCES


Konstantopoulos, P., Chapman, P. and Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accident Analysis & Prevention, 42, 827-834.


