

Producing a Commentary Slows Concurrent Hazard Perception Responses

**Angela H. Young
Peter Chapman
David Crundall**

This manuscript represents the final peer-reviewed and revised draft of the paper (26/3/14) as accepted for publication in the Journal of Experimental Psychology: Applied (JEPA). This has been added to the Nottingham Trent University I-Rep (internal repository) following guidelines available at <http://www.apa.org/pubs/authors/posting.aspx>.

The American Psychological Association retains copyright to all articles published in their journals. When this paper is published it will be available to download in APA format from <http://www.apa.org/pubs/journals/xap/index.aspx> for those with a journal subscription or who wish to purchase an individual copy. APA guidelines require the following statement to be included with any post-print versions: "This article may not exactly replicate the final version published in the APA journal. It is not the copy of record."

Running head: Producing a commentary slows hazard perception

Producing a Commentary Slows Concurrent Hazard Perception Responses.

Angela H. Young

Peter Chapman

David Crundall

University of Nottingham

Author Note

Angela H. Young, School of Psychology, University of Nottingham; Peter Chapman, School of Psychology, University of Nottingham; David Crundall, School of Psychology, University of Nottingham.

David Crundall is now at Division of Psychology, School of Social Sciences, Nottingham Trent University.

Our thanks go to Mike Swift for his demonstration of commentary driving and for sharing his experience of the use of commentary. Thanks also to Laura Thomson for her assistance in data collection.

Correspondence concerning this article should be addressed to Angela Young, School of Psychology, University Park Campus, University of Nottingham, Nottingham, NG7 2RD.

E-mail: lp xay@nottingham.ac.uk.

Key words: Commentary training; hazard perception; eye movements.

Abstract

Commentary driver training involves teaching drivers how to verbally acknowledge their perceptual and cognitive processes while driving, and has been shown to improve performance in driving-related tasks. However, those studies demonstrating benefits of commentary training have not done so under conditions of live commentary, which is the typical protocol used with advanced drivers. In the current study we present the results of two experiments that show that producing a commentary can actually slow responses to hazards on a concurrent hazard perception task. In Experiment 1 participants producing a live commentary showed significantly longer hazard response times than an untrained, silent, control group. In Experiment 2 a shorter, clipped commentary was introduced to attempt to reduce the demands placed upon participants. However, both the clipped and full commentary conditions showed reduced accuracy and longer response times, relative to a silent condition, and no difference was observed between the two types of commentary. Analysis of eye movements in both experiments revealed that fixation durations were shorter when a commentary was produced, but time to first fixate the hazard was not affected. This suggests that commentaries encourage more active interrogation of the visual scene, but that this can be detrimental to performance in average drivers.

Commentary driving is a training technique, currently used with advanced drivers and emergency services rapid response drivers (Coyne, 1997; Sharp, 1997). When used in the field, it involves training the driver to produce their own continuous verbal running commentary, while driving, on what they are doing (e.g. changing gear), what they can see (e.g. a pedestrian by the road), what might happen (e.g. the pedestrian may walk out into the road) and what they intend to do (e.g. slow down, move further out into the road) (Gregersen, 1994). This is expected to improve visual scanning and interpretation of the visual scene (Marek & Sten, 1977). There is mounting evidence that training novice drivers in commentary driving techniques can lead to improvements in their hazard perception (HP) and simulated driving performance.

Crundall, Andrews, van Loon, and Chapman (2010) had learner drivers complete a route in a driving simulator that contained several hazards. Half of the learner drivers then went on to be trained in commentary driving by their usual instructor. When the learners returned to the lab and completed the hazard route a second time, those trained in commentary had fewer crashes, slowed sooner on the approach to hazards and applied pressure to the brakes sooner than the control group. Benefits of commentary practice or exposure have also been observed with relatively little training. Listening to or producing a commentary while watching driving video clips has been shown to lead to improvements in subsequent HP test scores (Isler, Starkey, & Williamson, 2009; McKenna, Horswill, & Alexander, 2006; Wetton, Hill, & Horswill, 2013) and reduced risk taking on a gap acceptance task (McKenna et al., 2006). It should be noted however that, while some studies required drivers to produce commentaries in the training phase, the subsequent assessment phase was always held in silence.

The evidence so far suggests that commentary training could be a safe and economical way of improving drivers' HP skills and road safety. While training in the

laboratory might involve watching driving video clips, in practice learning to produce a commentary involves giving a live on-road commentary while controlling the vehicle, with or without supervision. Resources are readily available that allow interested drivers to teach themselves how to produce a commentary that can be practised on the road (Gilbert, 2007). However, as already noted, there has been no investigation into the effects of producing a driving commentary on concurrent HP or driving performance.

The importance of investigating the effect of a concurrent driving commentary is clear. While the evidence suggests that a trained commentary driver can negotiate hazards more effectively than an untrained driver, producing a concurrent commentary could have different effects.

A concurrent commentary could be most simply viewed as a secondary task, which could interfere with the primary task of controlling the vehicle or responding to hazards. The most well-documented example of a detrimental secondary verbal task is the use of mobile telephones (for a review see Caird, Willness, Steel & Scialfa, 2008), whether they are hand held or hands free (Redelmeier & Tibshirani, 1997). A detriment is also observed where drivers carry out other verbal tasks, such as the Paced Auditory Serial Addition Test (PASAT), in which drivers hear a number at pre-recorded intervals and are asked to add each digit to the digit preceding it and respond verbally (Rizzo et al., 2004; Thompson et al., 2012).

In both of these detrimental verbal tasks the driver processes an auditory input and produces a verbal output, with both the input and output being unrelated to the driving task. In the case of commentary, although a verbal response is required, there is no requirement for auditory processing. In fact, visual attention must be directed to the driving scene in order to process elements of the scene for verbal report.

A more comparable secondary task is concurrent verbalisation. Concurrent verbalisation is a technique intended to create a verbal record of what attracts the driver's attention and what they are looking at, without affecting their behaviour (Cole & Hughes, 1984; Hughes & Cole, 1986a, 1986b).

Crundall and Underwood (1997) investigated the effect of concurrent verbalisation on manual responses and eye movements during a HP test. The findings showed no effect of verbalisation on speed of hazard responses, or on fixation durations or variance in fixation locations. This provides some suggestion that a secondary verbal task need not interfere with hazard responses or search strategy when it does not involve redirecting attention away from the driving scene.

However, commentary production differs from concurrent verbalisation in important ways that could lead to distraction. In a concurrent verbalisation participants are asked to report what they see, leaving natural pauses when there is nothing further to report. In this case the visual behaviour of the participant determines the content of the verbal report. In a driving commentary the participant is additionally asked to make predictions about what might happen next, decide and report what action they would take to avoid developing hazards and not leave pauses. Hughes and Cole (1986a) suggested that producing a continuous verbalisation (without pauses) may lead participants to pay attention to things that they would normally ignore. The need to identify potentially hazardous situations, make predictions about the future behaviour of other road users and plan a response for verbal report might also be expected to lead to important changes in patterns of visual search and cognitive processes. Since these behaviours are vital for early hazard detection, the addition of these elements might be expected to change what the driver looks at in such a way as to improve hazard responses. This would, however, rely on the time course of processing for commentary and hazard perception to be the same. If attention was directed away from a

potential hazard source in order to locate the next item to comment on before a developing hazard had been detected, for example, hazard perception might be delayed.

Video based HP tests can be used to differentiate between novice and expert drivers (Borowsky, Oron-Gilad, & Parmet, 2009; Borowsky, Shinar, & Oron-Gilad, 2010; Horswill & McKenna, 2004; Isler et al., 2009; Scialfa et al., 2011) and have been shown to be predictive of crash risk (Wells, Tong, Sexton, Grayson, & Jones, 2008). They are safe and efficient to conduct, making them an ideal tool to test the effect of commentary on hazard perception.

While hazard responses can reveal the effect of commentary on speed and accuracy of hazard detection, further insight into the effect of commentary on visual search behaviour can be gained by observing eye movements. A number of differences in eye movements have been observed between novice and experienced drivers. Novice drivers have been shown to use less flexible search strategies, have a narrower spread of search and longer fixation durations than more experienced drivers (Chapman & Underwood, 1998; Crundall & Underwood, 1998; Mourant & Rockwell, 1972). Although the latter finding has not always been present (Underwood, Chapman, Bowden, & Crundall, 2002), it is believed that the shorter fixation durations of experienced drivers are indicative of an ability to process driving-related stimuli more quickly than novice drivers, due to increased experience. Novice drivers have also been shown to fail to look at locations that are relevant for the assessment of risk (Pradhan et al., 2005)

To date there has been no investigation of the effect of either training in commentary or producing a live driving commentary on eye movements. If producing a commentary improves visual search strategy, we first might expect this to be reflected in earlier or more frequent fixations on the developing hazard. We might also expect eye movements to become more like those of experienced drivers, with shorter fixations and a wider spread of

search. Alternatively, a concurrent commentary might mean that items being commented on capture attention for longer than usual, leading to longer fixations and a less efficient visual search, leading to later detection of the developing hazard.

In the current study we investigate the effect of producing a commentary on concurrent hazard perception skills and eye movements. In Experiment 1 we look at the effect of commentary production during a HP test on the frequency, accuracy and timing of hazard responses. We also look at differences in general eye movements (number of fixations and mean fixation duration) and eye movements to the hazard causing object (time and duration of first fixation, total fixation time). We expect fixation durations to be affected in the commentary condition, either becoming more like those of more experienced drivers or with longer fixations as commentary distracts attention from the search for hazards. In Experiment 2 we investigate whether limiting the number of words per utterance while producing a commentary might reduce demands and affect hazard responses and eye movements differently to a full commentary. Participants completed a number of hazard clips in silence, while giving a full commentary or while giving a clipped commentary, in which they were asked to keep each utterance to one or two words.

Experiment 1

In Experiment 1 we used a bespoke training video to train half of the participants in how to produce a driving commentary. Participants then practised their commentary over a practice hazard perception (HP) clip, before producing a commentary during a 13 clip HP test. The performance of these participants was compared to that of a control group who watched the same training video material without the accompanying audio, and completed the HP test in silence.

Eye movements were recorded and analysed to establish whether producing a commentary may lead to an improved search strategy or an earlier first fixation upon hazards.

Method

Participants Fifty-three students between 18 and 37 years of age ($M = 21$, $SD = 3$, 24 male) took part, with 27 randomly allocated to the control group and 26 to the commentary group. Participants were recruited as part of a larger study that began with the experiment reported here and was followed by a simulator study, which is not discussed further here. All participants were English speakers, had normal or corrected-to-normal vision, and had held a full UK driving licence for 1-19 years ($M = 3.2$, $SD = 2.9$).

Stimuli and apparatus Participants viewed a 30 minute bespoke training video consisting of footage of a drive around Nottingham (UK) filmed from the perspective of the driver. A 1080i full HD Sony video camera was used, with a Carl Zeiss Vario-Sonnar T* lens. The camera was mounted on the inside of the windscreen using a Manfrotto suction cup camera mount. Audio was recorded using a Sennheiser Ew 100 G2 radio microphone with a tie mic. The drive was undertaken by a professional Approved Driving Instructor (ADI), who had been trained in commentary driving and was himself an experienced instructor of commentary driving. Road and traffic conditions included suburban and urban roads and dual carriageways, incorporating Nottingham city centre, as well as quieter residential areas. The instructor provided a commentary throughout the 30 minute video, though it was possible to present the video with or without the accompanying audio. Sections of the training video can be viewed at www.lifelongdriving.org.

The HP test consisted of 13 video clips, ranging in length between 19.9 and 54.7 seconds ($M = 37.3$, $SD = 11.8$). These were filmed using a similar set up to that of the training video. Twelve of the clips contained one predefined hazard, while one clip contained two predefined hazards. For the purposes of this study a hazard was defined as any object or

event that might require you to suddenly change speed or change your position in the road in order to avoid a collision.

Button press responses that fell between the hazard onset and offset times were counted as correct responses to the predefined hazard and response times were recorded relative to the hazard onset time. The hazard onset time was initially determined *a priori* as the point at which a driver could predict that the occurrence of a hazard was unavoidable. However, in an earlier experiment using the same clips, participants had been required to pause the clip and verbally describe the hazard. The results from this experiment demonstrated that, for some of the clips, it was possible to correctly respond to the predefined hazard before the *a priori* hazard onset time. Where this was the case the hazard onset time in this experiment was adjusted to reflect this. Hazard onset times ranged from 9.7 to 45.6 seconds ($M=25.4$, $SD=11.1$). Hazard offset times were derived from the same earlier experiment, using an offset time of 3 standard deviations above the mean response time for each hazard.

The hazard clips were presented on a 22" flat screen Dell monitor at a resolution of 1680x1050, using SensoMotoric Instruments (SMI) Experiment Centre 3.0. Eye movements were recorded using an SMI iViewX™ Remote Eye tracking Device. The system was used at a temporal resolution of 60hz and a spatial resolution of 0.03°. The eye tracker is mounted at the bottom of the monitor and tracks gaze position of the participant at a viewing distance of around 60cm, without the need for a head restraint or chin rest. An Olympus WS-320M digital voice recorder was used to record commentary.

Design A between subjects design was employed, with two groups. Participants were randomly allocated to either the control group, who had no exposure to commentary, or the commentary group, who were trained in how to produce a commentary and were then subsequently required to give a commentary during the HP test. Training consisted of

watching a 30 minute video of commentary provided by an Approved Driving Instructor while driving around Nottingham. While watching this training video, all participants were asked to press a key for any perceived hazards. Control participants watched this video without sound and so received no commentary training, but were still required to press a key for any hazards in order to maintain engagement with the video.

Behavioural data for analysis were derived from the HP test and included frequency of hazard responses, number of correctly identified predefined hazards and response times. For analysis of number of correctly identified hazards and reaction times all cells in which the response count was more than 3SDs above the mean number of responses for the clip were counted as incorrect, as the participant had not effectively discriminated between the hazardous situation and non-hazardous parts of the clip.

For analysis of response times, trials in which no correct response was made were replaced by the maximum possible correct response time for that hazard. All analyses were repeated with no correct response treated as missing data. Where the two analyses differ it will be reported in the results.

Gaze coordinates were recorded at a rate of 60Hz. For analysis of eye movement data any consecutive set of gaze location coordinates with a dispersion of less than 100 pixels for at least 80ms were classified as a fixation. The intervening samples were classified as saccades or, if no data were recorded, as blinks. Number of fixations and mean fixation durations were analysed. In order to allow an analysis of attention to the hazard itself, each hazardous target (e.g. a vehicle or pedestrian defined as the *a priori* hazard) was used as an area of interest (AOI). The AOI encompassed the hazardous object, beginning at the point where the object first entered the scene, and following it until it exited the scene or the hazard window ended, whichever was the earlier. For each AOI the time at which the gaze first

entered the area, the number of fixations within the AOI area, the duration of the first fixation and the sum of all fixation durations (total dwell time) in the AOI were analysed.

Procedure After completing a short demographic questionnaire, participants were asked to face the monitor in a position that they could comfortably maintain for the following 30 minutes. Adjustments were then made to the position of the participant's chair and the monitor (with eye tracker attached) in order to place the eyes in approximately the centre of the eye tracker camera view.

All participants wore headphones for the duration of the training video and received onscreen instructions. The eye tracker was calibrated and the participants were instructed to watch the training video and press a key whenever they saw a hazard. A hazard was described as any object or event that might require you to suddenly change speed or change your position in the road in order to avoid a collision.

The experiment reported here forms part of the training procedure for an experiment involving a driving simulator, so both groups were told that they would later complete a HP test and drive in the simulator. The control group were told that they should pay attention to the video, as this was the only training that they would receive in how to drive safely. The commentary group were told that they may later be asked to provide their own commentary while carrying out the HP test and driving in the simulator and that they should pay attention to the video, as this would be the only training that they would receive in how to produce a commentary. The control group watched the video with no audio, while the commentary group watched the video with the accompanying commentary. To ensure that control participants remained engaged in the video, all participants were asked to press the zero key whenever they saw a hazard.

After completing the training video participants were encouraged to take a short break, before they resumed their position and the eye tracker was recalibrated. All

participants were advised that they would view 13 driving clips, 12 of which would contain one hazard and one of which would contain two hazards. Participants were asked to watch the clips and press the zero key as quickly as possible when they saw a hazard. They were also advised that they could press zero several times, but if they pressed too many times in a single clip their data would be rejected. The commentary group were asked to provide a verbal commentary over the clips, which was recorded. After completing a practice clip, the participants completed the HP test.

After the experiment, participants went on to use a driving simulator for a separate experiment.

Results

Across all participants, there were five instances where a participant pressed the wrong button to respond, resulting in an inadvertent termination of the clip. These responses were not included in the subsequent analyses (0.7% of cells).

The length of each HP clip varied, as did the visual properties of the clip and the length of the hazard window during which responses were treated as correct. This meant that some clips could have a greater effect on group means than others and missing data for a particular clip could lead to an over or underestimation of participant performance. For example, if a participant gives no correct response on a clip where the mean response time is longer than other clips, so have no response time for that clip, the mean response time for that participant will be artificially low. Equally, missing data for a clip with a particularly short mean response time would make the participant's mean response time artificially high. To ensure that each clip contributes equally to participant means the data for each clip were z-scored using the mean for that clip, before being entered into analyses. For ease of

interpretation the means and standard deviations reported here have been converted back to the original units from z-scores, using the grand mean and standard deviation.

Behavioural Analyses There was no significant difference between commentary (M = 1.1, SE = 0.1) and control groups (M = 1.2, SE = 0.1) in the mean number of hazard responses per clip, [$t(51) = 0.7, d = .20$].

For analysis of response accuracy and reaction times all cells in which the response count was more than 3SDs above the mean for the clip were counted as incorrect (2.2%), as the participant had not effectively discriminated between the hazardous situation and non-hazardous parts of the clip.

An independent *t*-test on the percent of the predefined hazards that were correctly identified showed no significant difference between the commentary (M = 87.4%, SE = 1.2) and control (M = 89.6%, SE = 1.3) groups, [$t(51) = 1.2, d = .34$].

Importantly, hazard response times were significantly longer in the commentary group (M = 1738ms, SE = 49.5) than the control group (M = 1557ms, SE = 36.3), $t(51) = 3.0, p < .005, d = .84$.

Eye Movement Analyses Data from two control and three commentary participants (9.4%) were excluded from the analysis due to missing eye movement data (missing data for more than 10% of the clip length on 4 or more clips).

There were slightly more fixations per clip in the commentary group (M = 118.2, SE = 1.1) than the control group (M = 114.6, SE = 1.4), but this difference did not reach statistical significance [$t(46) = 2.0, p = .054, d = .59$]. This was accompanied by a decrease in mean fixation duration, with the commentary group (M = 260.0ms, SE = 5.4) displaying significantly shorter fixations than the control (M = 281.2ms, SE = 6.7), $t(46) = 2.4, p < .05, d = .71$.

Every hazard was defined as an AOI. The AOI encompassed the hazardous object, beginning at the point where the object first entered the scene, and following it until it exited the scene or the hazard window ended, whichever was the earlier. Commentary did not significantly affect the time to first fixate on the hazard, [$t(46) = .64, d = .19$] ($M_{\text{COMMENTARY}} = 3412.3\text{ms}, SE = 65.9$; $M_{\text{CONTROL}} = 3350.7\text{ms}, SE = 49.8$), or the number of fixations on the hazard, [$t(46) = 1.36, d = .40$] ($M_{\text{COMMENTARY}} = 8.40, SE = 0.10$; $M_{\text{CONTROL}} = 8.17, SE = 0.10$). Although mean first fixation durations were slightly shorter for commentary ($M = 335.4\text{ms}, SE = 5.7$) than control conditions ($M = 350.9\text{ms}, SE = 5.9$), this difference did not reach statistical significance [$t(46) = 1.99, p = .053, d = .59$].

Discussion

The aim of Experiment 1 was to discover whether an online commentary resulted in faster and more accurate HP responses, or whether the production of a commentary might degrade hazard perception performance. The results did not suggest any change in the accuracy of HP responses in the commentary group compared to the control group, but response times were slowed for those participants who had to give a commentary. This suggests that giving a commentary degrades HP performance.

One possible mechanism that could underlie the degrading effects of online commentary would be an increase in fixation durations: when making an utterance in regard to a fixated object, we hypothesised that the length of the fixation on that object may be extended to the duration of the utterance. For instance, if a parked vehicle normally attracts a fixation of 400 ms, but the duration of the utterance lasts 2000 ms (e.g. “I see a parked vehicle on the left side of the road”), the participant may remain fixated on the object until the utterance is finished. This was not the case however. In fact, fixation durations were shorter, suggesting ocular motor strategies more similar to those of experienced drivers

(Chapman & Underwood, 1998). Indeed this is the effect that many practitioners of commentary driving would predict, with drivers paying more attention to a greater variety of cues in the driving scene. Despite the shorter fixation durations, these ostensibly-improved measures of eye movements do not carry through into faster hazard responses. Why might this be the case?

It is assumed that highly experienced drivers make shorter fixations than novice drivers because they are able to process driving-related stimuli more efficiently. It remains possible however that the act of producing a commentary may evoke shorter fixations for a different reason than a reduction in the processing difficulty of the fixated object. Although an increase in spatial load can lead to longer fixation durations while driving (Recarte & Nunes, 2000), non-spatial secondary tasks have been associated with a decrease in fixation durations during driving (Nunes & Recarte, 2002), as has naturalistic conversation (McCarley et al., 2004).

Although eye tracking data show that the fixation does not remain on the object of commentary for the duration of the utterance, it is still possible that participants are processing certain objects in working memory after their eyes have moved to another location in the visual scene. If a verbally-reported object continues to demand cognitive resources until the end of the utterance, even if it is no longer currently fixated, this might explain the degraded hazard perception performance. There are no data in the current experiment that can be used to address this issue. However if one could artificially evoke shorter, clipped utterances, this should allow this hypothesis to be assessed. Accordingly, a second experiment was undertaken where a full commentary condition, and a control condition, were both compared to a clipped commentary condition. In this latter condition, participants received specific instructions to keep utterances short and to the point. If the act of producing a commentary is increasing the amount of time that a participant devotes to processing an

object (independently of how long they look at it for), then shorter utterances should reduce this secondary demand and improve HP performance.

One further change to the experimental design addressed the possibility that differences between the groups were caused by the training period rather than the act of producing a commentary. In Experiment 1 participants were either untrained and silent at test or trained and producing a commentary at test. In Experiment 2 a within subjects design is used and participants watch only a very short example of a clipped or full commentary before producing their own at test. Effects of commentary would therefore be expected to be caused by the act of producing a commentary, rather than watching an extended expert commentary.

Experiment 2

In Experiment 2 we investigated the effect of a clipped commentary on behavioural and eye movement performance during a HP task, relative to a silent control condition and a full commentary condition, using a within subjects design with limited training.

In the clipped commentary condition participants are asked to restrain each utterance to one or two words, while in the full commentary condition they are asked to use natural language. Restraining the length of each utterance allows us to investigate the effects of a limited commentary, which may be both less demanding in terms of formulating a verbal response, and would be expected to shorten the time that participants need to spend attending to or processing each item that they describe.

If reductions in fixation durations represent an improvement in visual search strategy, but the secondary task of producing the verbal commentary delays the button press response, we might expect faster responses in the clipped condition. It remains a possibility however

that any form of commentary increases the demands of verbally-reported objects, independently of the length of the utterance.

Method

Participants Thirty six students between 18 and 25 years of age ($M = 20$, $SD = 1.7$, 9 male) took part. All participants were English speakers, had normal or corrected-to-normal vision, and had held a full UK driving licence for 1-7 years ($M = 2.8$, $SD = 1.4$).

Stimuli and Apparatus An HP test was used, consisting of 29 hazard clips filmed from the perspective of the driver. Each clip ran for between 10 and 62 seconds ($M = 39$ seconds, $SD = 16$). Twenty eight of the clips contained one predefined hazard and one clip contained two predefined hazards. The clips were split into three sets, each containing 10 hazards (either 9 or 10 clips) and totalling between 6 minutes 31 seconds and 7 minutes 30 seconds of footage. Nine of the clips were the same as those used in Experiment 1. Hazard onsets and offsets were defined *a priori*. Onsets were defined as the point at which a driver could predict that the occurrence of a hazard was unavoidable, while offsets were determined as the point at which a collision would have occurred if the camera car had not taken evasive action by reducing speed or changing road position.

Clips were presented and behavioural and eye movement data collected using the same equipment as in Experiment 1. Commentaries were recorded directly onto a second computer using Audacity software.

Design A within subjects design was employed, with three conditions. In the control condition participants were asked to complete the HP test in silence. In the two commentary conditions participants were asked to give a verbal commentary on what they were looking at and thinking about, as if they were the driver while completing the HP test. In the clipped commentary condition participants were asked to constrain each utterance to one or two

words. For example, “car left” or “slowing”. In the full commentary condition they were not asked to constrain the number of words, but to use natural language to describe events. For example, “There’s a car coming from the side road on the left, it might pull out” or “There’s an oncoming bus ahead, I’m slowing to wait”.

Every participant took part in all three conditions, each with a different clip set. The association between condition and clip set, the order of the clip sets and the order of the conditions were counterbalanced. The order of the clips within each clip set was randomised.

No training video was used in this experiment. Prior to each condition participants were given instructions on the commentary to provide, watched an example clip accompanied by the appropriate style of commentary, as provided by a driving expert, and then practised their own commentary over a second video clip. In the control condition the example clip was played without any accompanying audio and the participant completed the practice hazard clip in silence.

The same behavioural and eye movement data were analysed as in Experiment 1. Audio recordings were also used to determine the number of words used in the two commentary conditions to refer to the hazard. Audio files were also used for an in-depth analysis of the commentary produced; these data will be presented as part of a separate paper.

Procedure After completing a short demographic questionnaire participants were instructed on how to complete the HP test. They were informed that they would complete three short HP tests, with each test consisting of 9 or 10 clips, each containing one or two hazards. They were asked to watch the clips and search for hazards. A hazard was described as any object or event that might require you to suddenly change speed or change your position in the road in order to avoid a collision. Participants were asked to press zero as quickly as possible when they saw a hazard and advised that they could press zero several times, but that if they pressed too many times in a single clip their data would be rejected.

They were also told that they would be required to give a verbal commentary in some of the tests, but would receive separate instructions before each test.

The first set of clips varied depending on the counterbalancing condition, as did the condition in which it was completed (control, clipped commentary or full commentary).

Prior to each set of clips the participant was instructed on the requirements of the condition (see design); specifically whether they should complete the HP test in silence, provide a clipped commentary or a full commentary out loud while completing the test. They watched an example clip in silence, accompanied by a recorded clipped commentary, or a full commentary, depending on the condition. After the example clip the participant completed a practice clip, in which they were silent, gave a clipped commentary or gave a full commentary, as appropriate.

The participant's seating position was checked for optimum eye tracking and the eye tracker was calibrated before the clip set began. This procedure was then repeated for the two remaining conditions.

Results

The behavioural and eye movement data from one participant were removed prior to analysis, as they made no correct response for 8 out of 10 clips in one condition (pressed the response key more than 3SDs above the clip mean), suggesting that they were not completing the HP task appropriately.

Data were analysed using 1 x 3 repeated measures ANOVAs with three levels (control, clipped commentary and full commentary). Where the assumption of sphericity was violated Greenhouse-Geisser adjustments were used. Where this was the case the p values reported are after Greenhouse-Geisser adjustment, however the original degrees of freedom are retained for clarity. As both commentary conditions differ from the silent control

condition in their requirement for a verbal response, significant main effects were explored using planned Helmert contrasts comparing the control condition with the combined mean of the commentary conditions, and then comparing the clipped commentary condition with the full commentary condition.

All data were z-scored, as in Experiment 1. For ease of interpretation the means and standard deviations presented here have been converted back to the original units from z-scores using the original grand mean and standard deviation.

To ensure that participants were following instructions and that utterances were indeed shorter in the clipped commentary, the number of words used to refer to the hazard in each clip during the commentary conditions was analysed. The audio recordings for two participants were not available, due to recording software error. A repeated measures *t*-test on the remaining data showed that hazard references involved significantly more words in the full commentary condition ($M = 5.9$, $SE = 0.3$) than in the clipped commentary condition ($M = 3.1$, $SE = 0.2$), $t(32) = 10.7$, $p < .001$.

Behavioural Analyses Analysis of response frequency showed a main effect of commentary on the number of responses per hazard, $F(2, 68) = 7.5$, $MSe = .872$, $p < .01$, $p\eta^2 = .180$. Significantly fewer responses were made in the commentary conditions ($M_{FULL} = 2.6$, $SE = 0.2$; $M_{CLIPPED} = 2.7$, $SE = 0.2$) than the control condition ($M_{CONTROL} = 3.0$, $SE = 0.2$), $F(1, 34) = 15.3$, $MSe = 2.351$, $p < .001$, $p\eta^2 = .310$. There was no significant difference in number of responses between the clipped and full commentary conditions [$F(1, 34) = 1.3$, $p\eta^2 = .038$].

For analysis of response accuracy and reaction times all cells in which the response count was more than 3SDs above the mean for the clip were counted as incorrect (1.8%), as the participant had not effectively discriminated between the hazardous situation and non-hazardous parts of the clip.

A one way ANOVA on percent correctly identified predefined hazards revealed a main effect of commentary, $F(2, 68) = 5.2$, $MSe = .616$, $p < .01$, $p\eta^2 = .134$, with more correct responses in the control condition ($M = 86.5\%$, $SE = 0.6$) than the commentary conditions, $F(1, 34) = 14.0$, $MSe = 1.594$, $p < .001$, $p\eta^2 = .291$, but no significant difference between the clipped ($M = 84.3\%$, $SE = 1.0$) and full ($M = 82.9\%$, $SE = 1.0$) commentary conditions, [$F(1, 34) = 1.1$, $p = .308$, $p\eta^2 = .030$].

Response time results are shown in Figure 1. There was a significant main effect of commentary on response times, $F(2, 68) = 28.9$, $MSe = 3.664$, $p < .001$, $p\eta^2 = .460$. Responses took longer in the commentary conditions than in the control condition ($M = 1867.0\text{ms}$, $SE = 29.2$), $F(1, 34) = 64.4$, $MSe = 10.981$, $p < .001$, $p\eta^2 = .655$, but, once again, there was no significant difference between the clipped ($M = 2156.6\text{ms}$, $SE = 39.6$) and full ($M = 2164.6\text{ms}$, $SD = 38.3$) commentary conditions, [$F(1, 34) = .055$, $p = .815$, $p\eta^2 = .002$].

Eye Movement Analyses All trials with missing data for more than 10% of the total clip length were removed before analysis (6.7%).

There was a small increase in number of fixations per clip in the commentary conditions, but this did not reach statistical significance ($M_{\text{FULL}} = 118.4$, $SE = 2.0$; $M_{\text{CLIPPED}} = 117.5$, $SE = 2.2$; $M_{\text{CONTROL}} = 114.0$, $SE = 2.3$) [$F(2, 68) = 2.2$, $p = .120$, $p\eta^2 = .061$]. However, there was a significant main effect of commentary on average fixation duration (Figure 2), $F(2, 68) = 9.2$, $MSe = 1.787$, $p < .001$, $p\eta^2 = .213$, with shorter fixations in the commentary conditions than the control condition ($M = 318.4\text{ms}$, $SE = 9.4$), $F(1, 34) = 12.8$, $MSe = 4.9$, $p < .001$, $p\eta^2 = .273$, but no difference between clipped ($M = 298.5\text{ms}$, $SE = 10.2$) and full ($M = 289.7\text{ms}$, $SE = 9.0$) commentary [$F(1, 34) = 2.4$, $p = .133$, $p\eta^2 = .065$].

As in Experiment 1, each hazardous target (e.g. a vehicle or pedestrian defined as the *a priori* hazard) was used as an area of interest (AOI), which began at the first point that the target entered the scene and ended when the target exited the scene, or at the end of the

hazard window, whichever was the earlier. The data from one clip were removed as only three participants fixated on the hazard. Commentary did not significantly affect the time to first fixate on the hazard, [$F(2, 68) = 0.3, p = .863, p\eta^2 = .010$], or the number of fixations on the hazard, [$F(2, 68) = 1.3, p = .327, p\eta^2 = .038$]. However, there was a main effect of commentary on first fixation duration, $F(2, 68) = 7.5, MSe = 1.105, p < .01, p\eta^2 = .180$, and the total time spent fixating the hazard, $F(2, 68) = 5.5, MSe = .835, p < .01, p\eta^2 = .140$, with a shorter first fixation in the commentary conditions ($M_{FULL} = 336.5\text{ms}, SE = 9.7; M_{CLIPPED} = 341.9\text{ms}, SE = 10.9$) than the control condition ($M = 379.3\text{ms}, SE = 10.1$), $F(1, 34) = 12.2, MSe = 3.273, p < .01, p\eta^2 = .265$ and less time spent looking at the hazard overall ($M_{FULL} = 1498.4, SE = 42.1; M_{CLIPPED} = 1519.7, SE = 41.9; M_{CONTROL} = 1646.7, SE = 35.5$), $F(1, 34) = 13.2, MSe = 2.461, p < .001, p\eta^2 = .280$. This is consistent with participants looking at the hazard as frequently in all conditions, but making shorter fixations throughout the clip in the commentary conditions. There was no significant difference between clipped and full commentary on any AOI measure.

Discussion

In Experiment 2 it has been shown that producing a commentary reduces frequency, accuracy and speed of response to hazards in a HP test, even when a within subjects design is used and only minimal instruction is provided. This suggests that differences are due to the act of producing a commentary, rather than the training process.

Participants successfully limited their commentary in the clipped commentary condition, producing utterances related to the predefined hazard that were approximately half the length of utterances in the full commentary condition. Despite this, there were no significant differences in the effects of a clipped or full commentary on either eye movements

or HP measures, eliminating the possibility that distraction is caused when items capture attention for the length of the associated utterance.

General Discussion

In Experiment 1 we found that a live commentary leads to slower hazard responses and shorter fixation durations. In Experiment 2 this finding was repeated in a within subjects design, without a significant training period. Additionally, we found that a clipped commentary, using only one or two words per utterance, had the same effects as a natural commentary, and was no less detrimental to the speed of hazard perception. The increased power of a within subjects design in Experiment 2 allowed us to detect a significant decrease in both frequency and accuracy of responses in the clipped and full commentary conditions relative to the silent control condition. This is consistent with an explanation either at the level of hazard detection or hazard response. A failure to detect a hazard would clearly lead to fewer responses. However, a failure to respond to hazards in an appropriate time frame may also lead to fewer responses if participants inhibit button press responses that are no longer appropriate to the portion of the clip being viewed.

Training in commentary driving has been shown to improve subsequent hazard perception performance (Isler, Starkey, & Williamson, 2009; McKenna, Horswill, & Alexander, 2006; Wetton, Hill, & Horswill, 2013). However, our results show that when a live commentary is produced at test hazard responses are delayed relative to the untrained, silent control. This suggests that producing a commentary acts as a distraction from the concurrent HP task, and that this distraction is sufficient to negate any benefit of training.

Dual task interference can occur at the level of perception, central processing or motor response. Interference in perceptual systems would be expected if both tasks required the same perceptual system to carry out different tasks. With regards to the visual system in

particular, two different displays, some distance apart, cannot be monitored with the same level of visual acuity as a single display (Wickens, 1980). This is well illustrated in the case of driving while using a telephone handset to dial a telephone number (Tornros & Bolling, 2005). In the case of producing a commentary and pressing a response key for hazards, however, the perceptual input should be the same; those elements of the visual scene that are hazard relevant. This would, however, rely on the content and time course of processing for commentary and hazard perception to be the same. As discussed, the additional demands of commentary (over and above verbally reporting what is being observed) could affect where and when the participant attends, and delay the perception of hazards. However, eye movement analyses revealed that there was no effect of commentary on time to first fixate the hazard or number of fixations on the hazard. Although the total fixation duration on the hazard was shorter in the commentary conditions, this was indicative of reduced fixation durations overall, rather than a specific reduction in fixations on the hazard. This suggests that the live commentary does not slow people in fixating the hazard, but it does affect their ability to respond. These results are consistent with a post-perceptual bottleneck, as drivers look at the developing hazard just as soon, but responses are delayed.

Many of the findings from dual task interference research are consistent with a central processing bottleneck (see Pashler & Johnston, 1998 for a review), in which central processing for the second task is delayed due to either competition for limited central resources (Kahneman, 1973; McLeod, 1977b), or central processes that can only be carried out for each task in serial (e.g. Posner, 1978; Welford, 1981). Evidence for this type of bottleneck has been shown during simulated driving using a verbal response to a visual or auditory stimulus and a braking task (Levy, Pashler, & Boer, 2006), in which verbal responses slowed braking in response to a visual stimulus. A central bottleneck account of dual task interference explains the detrimental effect of commentary on hazard perception, as

the additional demands of commentary over concurrent verbalisation would call upon additional central resources for making predictions about the future behaviour of other road users. Although this explanation would be problematic for commentary driving practitioners, it is possible that practice in commentary would make the task sufficiently automatic to reduce the requirement for central resources enough to avoid having an impact on hazard response. Although some cases of practice allowing the central bottleneck to be bypassed have been recorded in the laboratory (Ruthruff, Van Selst, Johnston & Remington, 2006) this requires the same task to be repeated many times. Since on-road conditions are constantly varying the likelihood of tasks that depend on these conditions becoming automated is reduced (Lien, Ruthruff & Johnston, 2006).

Another possible source of interference occurs at the level of response execution (Dejong, 1993; Keele, 1973). Response execution bottlenecks are typically observed when the same effector system is used in both tasks (McLeod, 1977a; Ulrich et al., 2006) but they have also been observed when tasks require a verbal and manual response (Bratzke et al., 2008), as is required in the case of simultaneous commentary and hazard responses. However, if a response execution bottleneck was responsible for the delay in the manual response to hazards in the current experiment we would expect this effect to be ameliorated in the clipped commentary condition of Experiment 2, where the verbal demands were reduced. Since this is not the case it seems unlikely that a response execution bottleneck could adequately explain the effect of commentary on concurrent hazard perception responses.

The fixation durations of typical drivers who produce a commentary are shorter than those of typical drivers in the control condition, just as the fixation durations of experienced drivers are shorter than those of novice drivers (Chapman & Underwood, 1998). Shorter fixations can allow more rapid scanning of the visual scene and are generally interpreted as

suggesting that the driver can process driving related stimuli more quickly than those with longer fixations. However, the associated increase in response time to hazards suggests that shorter fixations are not indicative of more rapid processing in this case. A secondary spatial loading task can lead to longer fixations during driving (Recarte & Nunes, 2000). However, non-spatial cognitive loading tasks, including autobiographical memory recall and currency conversion tasks, have been shown to reduce fixation durations while driving (Nunes & Recarte, 2002). More specifically, naturalistic conversation while driving has also been associated with a decrease in fixation duration (McCarley et al., 2004). In this case longer response times to hazards suggest that shorter fixations are caused by the secondary task of producing a verbal commentary due to increased verbal cognitive load, rather than the more rapid processing of the visual scene observed in expert drivers.

There is a wealth of research showing that conversing causes a distraction from driving that is detrimental to driving performance (for example, Caird, Willness, Steel & Scialfa, 2008) and giving verbal responses to non-driving relevant stimuli can also have a detrimental effect (Rizzo et al., 2004; Thompson et al., 2012), but it is not clear which elements of speech lead to this detriment. Differences between conversation and commentary production can give some insight into which elements of conversation are the best candidates for the cause of distraction. Conversation is a complex task involving listening and comprehension of speech, followed by formulation and production of a response. Producing a commentary does not involve listening or comprehension and in order to formulate a response attention must be directed to the road scene itself. Since producing a commentary leads to a detriment in hazard perception it seems unlikely that listening or comprehension are key components of the detriment observed during conversation. This is consistent with earlier research showing that listening to a radio broadcast or audiobook does not disrupt driving (Strayer & Johnston, 2001).

Concurrent verbalisation is a task that also involves speech based on the road scene that does not require listening or comprehension, however, unlike commentary, this does not produce any detriment in concurrent hazard detection (Crundall & Underwood, 1997).

Concurrent verbalisation is based on Ericsson and Simon's (1984) Think Aloud Protocol and used to create a verbal record of where the trainee attends. Participants are asked to say aloud what they are looking at, leaving pauses where no additional objects are attended to. In commentary driving trainees are similarly encouraged to say what they are looking at, but also to make predictions about the behaviour of other road users and developing hazards and to avoid pauses in the commentary. Since a detriment is not observed in concurrent verbalisation, our results suggest that distractions can occur if speech does not replicate the natural observations of the driver in real time, even when the content of speech is the road scene itself. Furthermore, this effect cannot be explained by a reduction in eye movements, as analysis of fixation durations shows that fixations are shorter under conditions of commentary.

Hazard perception can be considered as a specific instance of the broader concept of situation awareness (Endsley, 1995). Three levels of understanding are required for effective situational awareness; perception of elements in the current situation, comprehension of current situation and accurate prediction of the future status of the situation. Commentary driving might be most effective in forcing drivers to improve their predictions about the development of hazards. However, it is possible that when trainees attempt to produce their own commentary after little commentary exposure, they focus on the easier task of making observations about what they can see, rather than making predictions about what might happen. This could mean that exposure to commentary is beneficial to average drivers because of the *exposure* to predictive information, but rehearsing a commentary may initially have little benefit until trainees are able to keep up with both observational statements and the

predictive statements that might help to improve predictive abilities. After sufficient practice in commentary the trainee may eventually benefit further from improved hazard prediction, but this benefit may still only be observable when the trainee drives silently, due to the increased demands of commentary production outweighing the benefits of improved hazard prediction.

To help compensate for the increased load of producing a commentary trainees could potentially reduce their general driving speed. It has in fact been noted that new commentary drivers are tempted to reduce their speed so that they can produce an effective commentary (Gilbert, 2007). In this case the drivers are advised to attempt to continue driving at their normal speed and accept that they will not be able to comment on everything relevant at first. If drivers are able to do this it could reduce demands from the secondary task sufficiently to allow the primary task to be completed safely while the commentary is improved over time. However, in some cases a secondary task has been shown to interfere with a driving task even when participants are specifically told not to carry out the secondary task (Levy & Pashler, 2008). This raises questions about whether it is possible to entirely disregard the commentary task in dangerous driving situations, or whether the original intention to produce a commentary may still interfere with performance.

Despite evidence that commentary training can lead to improved hazard responses, our findings suggest that producing a live verbal commentary can actually impede hazard perception. Although there remains hope that well practiced driving responses may be less susceptible to distraction from live commentary production than a relatively novel button pressing response, the risk of distraction means that commentary practice should not currently be carried out on the road unsupervised.

References

- Borowsky, A., Oron-Gilad, T., & Parmet, Y. (2009). Age and skill differences in classifying hazardous traffic scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(4), 277-287. doi: 10.1016/j.trf.2009.02.001
- Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis and Prevention*, *42*(4), 1240-1249. doi: DOI 10.1016/j.aap.2010.02.001
- Bratzke, D., Ulrich, R., Rolke, B., Schroeter, H., Jentzsch, I., & Leuthold, H. (2008). Motor limitation in dual-task processing with different effectors. *Quarterly Journal of Experimental Psychology*, *61*(9), 1385-1399. doi: 10.1080/17470210701536856
- Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention*, *40*(4), 1282-1293.
- Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*(8), 951-964.
- Cole, B. L., & Hughes, P. K. (1984). A Field Trial of Attention and Search Conspicuity. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *26*(3), 299-313.
- Coyne, P. (1997). *Roadcraft: The Essential Police Driver's Handbook*.
- Crundall, D., Andrews, B., van Loon, E., & Chapman, P. (2010). Commentary training improves responsiveness to hazards in a driving simulator. *Accident Analysis and Prevention*, *42*(6), 2117-2124. doi: 10.1016/j.aap.2010.07.001
- Crundall, D., & Underwood, G. (1997). Concurrent verbalisation during drivers' visual search and hazard perception. In T. Rothengatter & E. Carbonell Vaya (Eds.), *Traffic and*

Transport Psychology: Theory and Application. North Holland: Elsevier Science Publishers B.V.

Crundall, D., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, *41*(4), 448-458. doi: 10.1080/001401398186937

Dejong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 965-980. doi: 10.1037/0096-1523.19.5.965

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32-64.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*. Massachusetts, USA: MIT Press.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1058-1079. doi: 10.1037/0096-1523.18.4.1058

Gilbert, C. (Writer). (2007). Ultimate driving craft [DVD].

Gregersen, N. P. (1994). Systematic cooperation between driving schools and parents in driver education, an experiment. *Accident Analysis and Prevention*, *26*(4), 453-461. doi: 10.1016/0001-4575(94)90036-1

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A Cognitive Approach to Situation Awareness: Theory and Application* (pp. 155-175). Aldershot, UK: Ashgate Publishing Limited.

- Hughes, P. K., & Cole, B. L. (1986a). Can the conspicuity of objects be predicted from laboratory experiments? *Ergonomics*, *29*(9), 1097-1111. doi: 10.1080/00140138608967226
- Hughes, P. K., & Cole, B. L. (1986b). What attracts attention when driving? *Ergonomics*, *29*(3), 377-391. doi: 10.1080/00140138608968272
- Isler, R. B., Starkey, N. J., & Williamson, A. R. (2009). Video-based road commentary training improves hazard perception of young drivers in a dual task. *Accident Analysis and Prevention*, *41*(3), 445-452. doi: 10.1016/j.aap.2008.12.016
- Kahneman, D. (1973). *Attention and Effort*. New York: Prentice Hall.
- Keele, S. W. (1973). *Attention and human performance*. California: Goodyear.
- Levy, J., & Pashler, H. (2008). Task prioritisation in multitasking during driving: Opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing. *Applied Cognitive Psychology*, *22*(4), 507-525. doi: 10.1002/acp.1378
- Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving - Is there any stopping the psychological refractory period? *Psychological Science*, *17*(3), 228-235. doi: 10.1111/j.1467-9280.2006.01690.x
- Lien, M.-C., Ruthruff, E., & Johnston, J. C. (2006). Attentional limitations in doing two tasks at once: The search for exceptions. *Current Directions in Psychological Science*, *15*, 89-93.
- Marek, J., & Sten, T. (1977). *Traffic environment and the driver: Driver behavior and training in international perspective*. Springfield: Charles C. Thomas.
- McCarley, J. S., Vais, M. J., Pringle, H., Kramer, A. F., Irwin, D. E., & Strayer, D. L. (2004). Conversation disrupts change detection in complex traffic scenes. *Human Factors*, *46*(3), 424-436. doi: 10.1518/hfes.46.3.424.50394

- McKenna, F. P., Horswill, M. S., & Alexander, J. L. (2006). Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology-Applied*, *12*(1), 1-10. doi: 10.1037/1076-898x.12.1.1
- McLeod, P. (1977a). Dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology*, *29*, 651-667. doi: 10.1080/14640747708400639
- McLeod, P. (1977b). Parallel processing and the psychological refractory period. *Acta Psychologica*, *41*(5), 381-396. doi: 10.1016/0001-6918(77)90016-6
- Mourant, R. R., & Rockwell, T. H. (1972). Strategies of Visual Search by Novice and Experienced Drivers. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *14*(4), 325-335. doi: 10.1177/001872087201400405
- Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 133-144. doi: [http://dx.doi.org/10.1016/S1369-8478\(02\)00012-8](http://dx.doi.org/10.1016/S1369-8478(02)00012-8)
- Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155 - 189). Hove, England: Psychology Press.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.
- Pradhan, A., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. *Human Factors*, *47*(4), 840-852.
- Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology-Applied*, *6*(1), 31-43. doi: 10.1037//0278-7393.6.1.31
- Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, *336*(7), 453-458.

Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. W. (2006). How does practice reduce dual-task interference: integration, automatization, or just stage-shortening?

Psychological Research, 70(2), 125-42.

Scialfa, C. T., Deschênes, M. C., FERENCE, J., Boone, J., Horswill, M. S., & Wetton, M.

(2011). A hazard perception test for novice drivers. *Accident Analysis & Prevention*, 43(1), 204-208. doi: <http://dx.doi.org/10.1016/j.aap.2010.08.010>

Sharp, G. (1997). *Human Aspects of Police Driving*. Clackmannanshire, UK Scottish Police College.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science*, 12, 462-466.

Thompson, K. R., Johnson, A. M., Emerson, J. L., Dawson, J. D., Boer, E. R., & Rizzo, M.

(2012). Distracted driving in elderly and middle-aged drivers. *Accident Analysis & Prevention*, 45(0), 711-717. doi: 10.1016/j.aap.2011.09.040

Tornros, J. E. B., & Bolling, A. K. (2005). Mobile phone use - Effects of handheld and handsfree phones on driving performance. *Accident Analysis and Prevention*, 37(5), 902-909. doi: 10.1016/j.aap.2005.04.007

Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving: skill and awareness during inspection of the scene. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(2), 87-97. doi: 10.1016/s1369-8478(02)00008-6

Welford, A. T. (1981). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction time*. New York: Academic Press.

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A study of learner and new drivers*. London: Department for Transport. Retrieved from

<http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme2/cohort2/cohortiiifindings.pdf>.

Wetton, M. A., Hill, A., & Horswill, M. S. (2013). Are what happens next exercises and self-generated commentaries useful additions to hazard perception training for novice drivers? , *54C*, 57-66. doi: 10.1016/j.aap.2013.02.013

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and Performance VIII*. Englewood Cliffs, NJ: Erlbaum

Figure 1

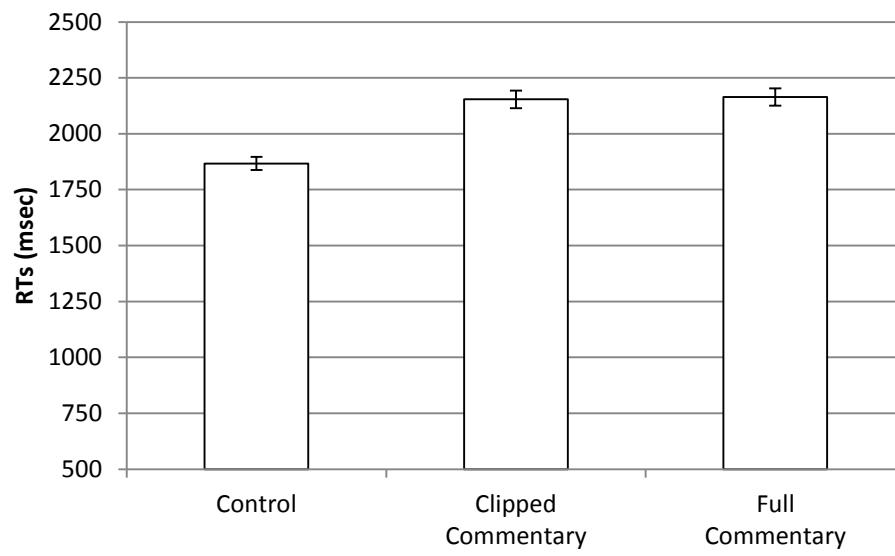


Figure 2

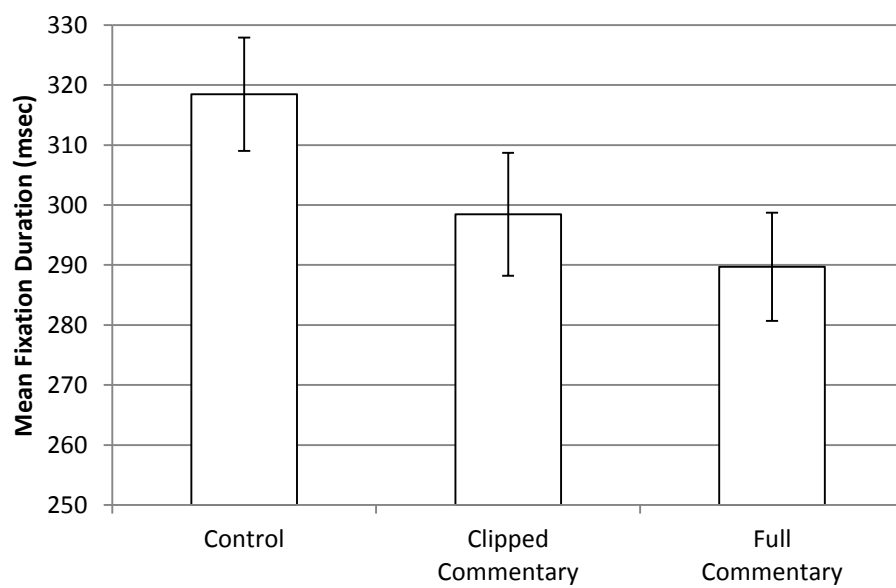


Figure Captions

Figure 1. Time to respond to the hazard (ms), measured from the hazard onset, in the control, clipped commentary and full commentary conditions. Error bars indicate one standard error above and below the mean.

Figure 2. Mean fixation durations (ms) in the control, clipped commentary and full commentary conditions. Error bars indicate one standard error above and below the mean.