Running head: commentary training

DOI: 10.1016/j.aap.2017.01.007

Commentary driver training: Effects of commentary exposure, practice and production on hazard perception and eye movements

Angela H. Young[1]

David Crundall[1]

Peter Chapman[2]

[1]Nottingham Trent University

[2]University of Nottingham

Author Note

Angela H. Young, Division of Psychology, School of Social Sciences, Nottingham Trent University, Nottingham, UK; David Crundall, Division of Psychology, School of Social Sciences, Nottingham Trent University, Nottingham, UK; Peter Chapman, School of Psychology, University of Nottingham, University Park, Nottingham, UK. Our thanks go to Mike Swift for his demonstration of commentary driving and for sharing his experience of the use of commentary.

Correspondence concerning this article should be addressed to Angela Young, Division of Psychology, School of Social Sciences, Nottingham Trent University, Nottingham, UK. E-mail: angela.young@ntu.ac.uk.

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

Abstract

Commentary driving typically involves being trained in how to produce a verbal running commentary about what you can see, what you are doing, what might happen and what action you will take to avoid potential hazards, while driving. Although video-based commentary training has been associated with subsequent hazard perception improvements, it can have a negative impact on hazard perception when a live commentary is produced at test (Young, Chapman, & Crundall, 2014). In the current study we use balanced training and testing blocks to isolate the effects of commentary exposure, production of a commentary with and without practice, and learning from earlier self-generation of commentary on behavioural and eye movement measures. Importantly, both commentary exposed and unexposed groups gave hazard perception responses during the commentary video, ensuring that the unexposed control group remained engaged in the procedure throughout. Results show that producing a live commentary is detrimental to concurrent hazard perception, even after practice, and does not enhance any later effect of commentary exposure. Although commentary exposure led to an initial increase in the accuracy of hazard perception responses, this effect was limited to the first occasion of testing, and showed no later benefits relative to engaged hazard exposure.

1. Introduction

Commentary driving involves producing a continuous verbal running commentary on what you are doing while driving, (e.g. changing gear), what you can see (e.g. a pedestrian by the road), what might happen (e.g. the pedestrian may walk out into the road) and what you intend to do (e.g. slow down, move further out into the road) (Gregersen, 1994). It is a training technique that is used to train advanced drivers and emergency services' rapid response drivers (Coyne, 1997; DIA, 2014; IAM, 2014; Sharp, 1997) and is expected to improve visual scanning and interpretation of the visual scene (Marek & Sten, 1977). In New Zealand new drivers are required to produce a commentary during their Full Licence Test (New Zealand Transport Agency, 2016), encouraging novice drivers to practice a commentary on the road.

Given that commentary involves the verbal prediction of hazards, it is reasonable to assume that the main benefit of commentary is derived from an improvement in hazard perception. Hazard perception is a driving-specific skill that is reliably correlated with collision risk (Wells, Tong, Sexton, Grayson, & Jones, 2008) and video-based hazard perception tests have been shown to differentiate between novice and experienced drivers (Borowsky, Oron-Gilad, & Parmet, 2009; Borowsky, Shinar, & Oron-Gilad, 2010; Horswill & McKenna, 2004; Isler, Starkey, & Williamson, 2009; Scialfa et al., 2011). This means that improving hazard perception skills might also reduce collision risk.

There is some evidence that training novice drivers in commentary driving techniques can lead to improvements in hazard responses in simulated driving (Crundall, Andrews, van Loon, & Chapman, 2010), and that simple video-based commentary exposure and/or practice can improve hazard perception (HP) responses (Horswill, Taylor, Newnam, Wetton, & Hill, 2013; Isler et al., 2009; McKenna, Horswill, & Alexander, 2006; Spolander, 1990; Wetton, Hill, & Horswill, 2013). Video-based commentary driver training is economical and easily accessible to drivers looking to improve their hazard perception skills; making it a desirable training method, providing it is effective. Indeed, apparent benefits have been observed in some studies with exposure to commentary videos of only eight minutes in length (Castro et al., 2016).

Although typical commentary driver training would involve both *exposure* to an expert commentary and *practice* at producing a live driving commentary (Crundall et al., 2010; Horswill et al., 2013; Spolander, 1990), it is not clear whether both elements are key to realising any benefits of commentary training. Benefits have been observed when drivers are trained via exposure to commentary, without ever producing a commentary of their own (Castro et al., 2016, McKenna et al., 2006; Wetton et al., 2013). However, improvements in hazard perception have also been observed when trainees produced their own commentary, without ever being specifically trained in how to produce a commentary (Isler et al., 2009). This might suggest that both commentary exposure and practising commentary production yield separate benefits for later silent hazard perception. In an attempt to disentangle these two effects Wetton et al. (2013) tested the hazard perception abilities of participants after either video-based commentary exposure alone, or a combination of exposure and practising a silent self-generated commentary. Although no additional benefit of the self-generated commentaries was observed, the silent nature of the produced commentaries made it impossible to determine whether the commentary protocol was correctly adhered to.

The question therefore remains whether there are separable effects of practising commentary production and being exposed to expert commentary, on subsequent hazard perception skill. In the current study we decompose training into video-based commentary exposure and separate commentary practise. If producing a commentary out-loud has benefits over and above those of video-based commentary exposure we would expect those participants who watch a commentary video and produce a commentary to perform better in subsequent silent testing, than those who watch the same video but do not practise producing a commentary.

Although we may not yet know whether commentary production is necessary in order to derive the greatest benefits from commentary driving, since current commentary training typically involves production of a commentary on the road, it is important to understand the effects of live commentary on concurrent hazard perception. Not only do trainees produce on-road commentaries with supervision, they are also able to use commercially available DVDs that demonstrate an expert commentary and suggest that the viewer carry out their own on-road commentaries (Gilbert, 2007). Young et al. (2014) found that when commentary exposure is followed by production of a driving commentary during a hazard perception test, hazard perception response times are significantly longer than in a control group that was neither exposed to commentary nor producing a commentary at test. This has important implications for commentary driver training, suggesting that trainees should not be encouraged to perform their own commentary on the road without expert supervision.

In Young et al. (2014) the observed slowing of hazard perception responses occurred when trainees were exposed to commentary using a video-based commentary example, and then produced their own commentary at test. It seems unlikely that commentary exposure caused the observed detriment, given that other research has shown benefits of commentary exposure, so the commentary production element is the prime candidate for having caused interference. Indeed, this fits well with the dual-task literature, which typically shows that when two tasks are carried out concurrently, performance on one or both tasks is inferior to either task when performed alone (for a review see Pashler, 1994). However, the extent of interference depends on the precise timing of the task demands and, in particular, whether both tasks require access to a limited capacity resource at the same time (see Pashler, 1994; or Wickens, 1980). For this reason, it is not clear whether interference in the earlier experiment was caused by talking about the road scene generally, while carrying out hazard perception, or is limited to the more active task of commentary driving. Talking while driving is not necessarily detrimental to performance in all cases. For example, although talking on a mobile telephone is known to cause distraction (for a review see Caird, Willness, Steel, & Scialfa, 2008) talking to a passenger is not always associated with poorer performance (Crundall, Bains, Chapman, & Underwood, 2005). Additionally, Hughes and Cole (1986) have used a technique known as concurrent verbalisation to gain insight into what viewers are doing when visually inspecting a driving scene. Like commentary driving, this involves verbalising what they are looking at, though without being required to give a continuous commentary or make predictions about the future status of objects in the scene. Hughes and Cole (1986) suggest that a verbal description of what is attended to should not affect processing demands. This idea was supported when Crundall and Underwood (1997) found that producing a concurrent verbalisation while carrying out a hazard perception test did not affect either behavioural responses or eye movements.

Since commentary driving involves talking about the contents of the visual scene, which should be attended to for hazard perception, one might expect the demands to be lower than other types of conversation but a detriment is still observed (Young et al., 2014). In order to establish whether the detriment of a live commentary is specific to producing a commentary following an earlier example or caused by any speech about the visual scene while searching for hazards we can compare the effects of producing a commentary after exposure to those of producing a naïve commentary, based on only limited instruction without commentary exposure. This is because a naïve commentary is subject to fewer rules about what should be said and how, so would be expected to be less effortful. If the detriment is specific to commentary following exposure, we might paradoxically find that commentary exposed drivers' HP performance is worse than that of naïve commentary producers when both groups give a commentary simultaneously with the HP test.

When drivers are fully trained in commentary driving they are initially supervised while gaining experience in commentary production. Even when a driver learns commentary by watching a DVD it is recommended that she practise a commentary over driving clips before doing so on the road (Gilbert, 2007). If detrimental effects are only present in the first few attempts at commentary production the outlook for commentary as a training method would be much brighter than if live commentary is detrimental even after practice. For this reason, it is important to investigate how the effect of commentary production on hazard perception might change with increased commentary experience.

When benefits of video-based commentary exposure or full training have been shown relative to a control group, the control group has typically watched the same commentary video, but without any accompanying audio, so that they were exposed to the same visual materials but without any commentary exposure (Horswill, Kemala, Wetton, Scialfa, & Pachana, 2010; Horswill et al., 2013; Isler et al., 2009; McKenna et al., 2006; Poulsen, Horswill, Wetton, Hill, & Lim, 2010; Wallis & Horswill, 2007; Wetton et al., 2013). However, this silent material may not encourage participants to engage with the video or expect any benefits of training. This raises the possibility that the observed benefits of video-based commentary training could in fact be due, at least in part, to a placebo effect. In Young et al. (2014) participants in both the commentary exposed and control groups were asked to search for and respond to hazards during the commentary video, to ensure that *all* participants were engaged in watching the footage, not just those listening to commentary. Although this reduced the likelihood of placebo effects, the effect of commentary production was confounded with that of commentary exposure. In the current study placebo effects are addressed and commentary exposure and production effects are disentangled.

While behavioural data can give an indication of the effect of commentary exposure and practice on hazard perception speed and accuracy, eye movement data can give deeper insight into how commentary affects attention in the visual scene. There has been little research into the effect of commentary exposure or practice on visual search of the driving scene. Young et al. (2014) found that those exposed to commentary and producing a commentary at test had shorter fixation durations than those who were neither exposed to nor producing a commentary. If commentary *exposure* produces shorter fixation durations then this may give us hope that commentary exposure is encouraging more rapid processing of the content of fixations, as is seen in experienced drivers relative to novice drivers (Chapman & Underwood, 1998). However, it is also possible that the shorter fixation durations are due to the *production* of a commentary, as shorter fixation durations have been associated with increased verbal workload (McCarley et al., 2004). Since Young et al. (2014) did not test silent hazard perception ability after commentary exposure or practice, it was not possible to establish whether shortened fixation durations were due to the production of speech or an effect of commentary exposure.

In the current study eye movements are recorded in speaking and silent participants, with and without commentary exposure and practice. In this way, we can differentiate between the effects of speech, commentary exposure and commentary production on eye movements. If shorter fixation durations are caused by commentary exposure then we would expect shorter fixations in the commentary exposed groups, regardless of speech condition. If, however, shorter fixation durations are caused by an increase in verbal workload, we would expect to observe them only in the speaking conditions, regardless of commentary exposure or prior commentary practice.

Other differences in visual search strategy might be expected as an effect of exposure and practice. An earlier fixation on a developing hazard gives a driver additional time to respond to the hazard. When eye movements were recorded during a live commentary (Young et al., 2014) there was no observed effect of commentary on time to first fixate a hazard, or total fixation duration on the hazard. However, there has been no investigation of the effect of commentary exposure/practice on hazard fixations during subsequent silent hazard perception. If the advantages of commentary are only observable when the trainee is silent we may also observe earlier or more fixations on developing hazards in exposed silent conditions.

In the current study, we use multiple training and testing blocks to answer several important questions about commentary exposure, production and practice. Training and testing conditions are summarised in Table 1. The experiment begins with all participants watching a 30 minute video filmed from the perspective of the driver. The commentary *exposed* participants see this video with accompanying audio of an experienced commentary driver giving a commentary in time with the video. The *unexposed* group watch the same video with no audio, so have no commentary exposure. This is the exposure factor. Crucially, participants in both the exposed and unexposed groups pressed a key to respond to hazards throughout the commentary video to ensure participant engagement with the material.

Table 1. The conditions experienced during the commentary video and hazard perception (HP) blocks 1, 2 and 3. A 2 x 2 x 3 mixed design is shown. Each row represents a unique condition created by the two between-subjects factors of ‘exposure’ and ‘speech in blocks 1 & 2’. The experiment timeline shows whether participants watched the video with audible commentary or on mute, and whether they were required to speak in HP blocks 1, 2 and 3.

Groups		Experiment Timeline			
Exposure	Speech	Video	HP block 1	HP block 2	HP block 3
Exposed	**Silent**	Hear commentary	Silent	Silent	Silent
	Speaking	Hear commentary	Commentary	Commentary	Silent
Unexposed	**Silent**	Mute	Silent	Silent	Silent
	Speaking	Mute	Commentary	Commentary	Silent

Half of the *exposed* and *unexposed* groups go on to produce a self-generated commentary in two following blocks of hazard perception clips, while the remaining participants carry out the same tests in silence. This is the *speech* factor (*block 1 & 2 speaking/always silent*). In HP blocks 1 and 2, the commentary exposed speaking participants are asked to base their commentary on the earlier video, while the unexposed speaking participants, having had no exposure to commentary, are given a very brief description of commentary and asked to produce their own. In a third HP block all participants are silent.

Each block allows us to address a different question about the effects of commentary exposure and practice. In the first block we can ask whether simple commentary exposure has benefits for subsequent silent hazard perception, by looking at the exposed and unexposed silent participants. Additionally, by comparing the exposed and unexposed participants while they are speaking, we can establish whether the predicted detrimental effects of live commentary production are specific to an exposed commentary, or due to talking about the road scene more generally.

In block 2 we can establish whether the predicted detrimental effects of commentary production are limited to the first few attempts at commentary production, in which case we would expect some improvement in the second block, or whether the detriment persists even after a whole block of practice.

Finally, in the third block we can assess the relative impact of commentary exposure and commentary practice on silent hazard perception. In this block, all participants are silent, allowing us to compare those who have been exposed to commentary (the exposed-silent group) with those who have been both exposed to commentary and practised a commentary in earlier blocks (exposed-block 1 & 2 speaking) to establish whether there is an additional benefit of commentary practice over and above that of commentary exposure.

In addition to the comparisons above that will inform our understanding of exposure and practice on behavioural measures, this design also allows us to question whether reduced fixation durations observed in those who were exposed to and producing a live commentary in Young et al. (2014) were due to commentary exposure or commentary production. If shorter fixations are due to exposure, we would expect to observe them in exposed conditions regardless of whether participants are required to produce a commentary. If however they are induced by the mere effort of general speaking we would expect to see shorter fixations in the speaking conditions (regardless of exposure) in blocks 1 and 2 only, disappearing in block 3 when all participants are required to be silent . Finally, if they are specifically caused by the production of a commentary after exposure we would expect to see them in the exposed speaking group in blocks 1 and 2.

2. Method

2.1. Participants

Forty-eight participants were randomly allocated to one of the four between-subjects conditions. The data from two participants had to be removed, as they did not follow instructions in the hazard perception test. This left 12 participants in each of the unexposed conditions and 11 in each of the exposed conditions. These participants were between 18 and 41 years of age (M = 20, SD = 3.7, 11 male), had normal or corrected to normal vision, spoke good English and had held a full UK driving licence for between 1 and 23 years (M = 2.5, SD = 3.5). The four groups did not differ significantly in age, time since passed test or estimated weekly driving hours.

2.2. Stimuli and apparatus

Participants viewed a 30 minute bespoke commentary video, before completing a hazard perception test consisting of 29 hazard clips split into three sets, each containing 10 hazards (across either 9 or 10 clips) and totalling between 6 minutes 31 seconds and 7 minutes 30 seconds of footage. Environments were either suburban (62%), urban (21%) or on a multi-lane carriageway (17%). Eye movements were recorded using an SMI iViewX™ Remote Eye tracking Device. See Young et al. (2014) for further details of the commentary video, hazard perception test (Experiment 2) and eye tracking setup used.

2.3. Design

The study design is shown in Table 1. A 2 x 2 x 3 mixed design was used, with two between-subjects factors of commentary exposure (exposed/unexposed) and speech (block 1 & 2 speaking/always silent). The exposure factor was implemented by having participants watch the commentary video either with or without the accompanying commentary audio. The within-subjects factor of testing block refers to the three HP testing blocks, with participants either always silent or speaking in blocks 1 & 2 (depending on speech condition). All participants were silent in the third HP block.

All participants took part in each of the three within-subjects conditions in the same order, but the association between clip set and condition was counterbalanced and the order of clips within each clip set was randomised.

2.4. Procedure

All groups were asked to watch the 30 minute commentary video and were told that they would later complete an HP test. The unexposed groups watched the video with no audio, so were not exposed to commentary, while the exposed groups watched the video with the accompanying commentary. The exposed groups were told that they may later be asked to provide their own commentary while carrying out the HP test 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 unexposed groups 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. We felt it was important for the groups who were not exposed to the expert commentary to still believe that they were being trained in a driving-safety related skill, to ensure that the demand characteristics of the experiment were equated across all participants. To ensure that they remained engaged in the video, all participants were asked to press the zero key whenever they saw a hazard.

For the HP tests participants 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. They were also asked to press the zero key on the numerical keypad as quickly as possible when they saw a hazard and that they could press several times during a particular clip to record as many hazards as they thought were visible, but they were warned that if they pressed too many times, their data would be rejected.

Before the first HP block ‘always silent’ participants were given the instructions, “Please remain silent throughout the hazard perception test.” These instructions were repeated before each of the subsequent HP blocks. Those unexposed participants who were expected to produce a commentary in blocks 1 and 2 were told “In the first hazard perception test please give a verbal commentary. To give a verbal commentary you should say out loud: What you can see (e.g. vehicle ahead, bend in the road), what might happen (e.g. car might pull out ahead, pedestrian could step out) and what you would do as the driver (e.g. reduce speed, move to the left). You should speak continuously throughout the clip, but don’t forget to press the zero key when you see a hazard.” Exposed participants who were also expected to provide a commentary in blocks 1 and 2 were told to “Please use the commentary you have just watched to guide your own commentary”.

Both of the block 1 & 2 speaking groups were asked to continue giving a commentary through the second hazard perception test. All participants were asked to remain silent during the third hazard perception test.

2.5. Data Analysis

Behavioural data for the analyses were derived from the HP test and included percentage of predefined hazards correctly identified and response times. 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. 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, to avoid artificially low mean reaction times for those who fail to respond to some hazards.

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. The mean AOI size relative to the video frame was 2.85% and mean visible AOI time was 9.06 seconds, or 24% of the video clip length. 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. Time at which gaze entered the AOI was measured relative to the hazard onset. Since for some hazards the hazardous object was visible prior to the hazard onset (e.g. When the vehicle/pedestrian was not presenting any indication of becoming hazardous), negative values have been included in this analysis, to avoid unfairly penalising those who fixated the object as it entered the scene and then monitored its progress peripherally, perhaps only fixating it again at a point in the clip that might be later than someone who had not noticed the object entering the scene and whose attention was only drawn by the development of the hazard.

3. Results

Data were analysed using 2 x 2 x 3 mixed design ANOVAs. 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. Significant main effects of block are followed up with planned difference contrasts, to compare the first two blocks (in which some participants speak) with the third HP block (in which all participants are silent), and to compare the first two HP blocks with each other.

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 is missing eye data on a clip where the mean time before the first fixation on the hazard is longer than in other clips, so has no 'time to first hazard fixation' data for that clip, the average time to fixate all hazards will be artificially low for that participant. 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 errors reported here have been converted back to the original units from z-scores, using the grand mean and standard deviation.

3.1. Commentary video

All participants watched the commentary video, either with or without audio in the exposed and unexposed groups respectively. Hazard responses, number of fixations, fixation count and vertical and horizontal variance in fixation locations were recorded while participants watched the video. Independent *t* tests on each of these measures showed no significant difference between the exposed and unexposed groups. Speech in blocks 1&2 was not included in analyses of training video data, as this manipulation was not implemented until after participants watched the training video.

3.2. Hazard perception responses

3.2.1 Response accuracy.

Responses are treated as incorrect when either no response is made during the hazard window or a participant makes a large number of responses, suggesting that have adopted too generous a criterion for the identification of hazards. 'Too many responses' was quantified as more than 3SDs above the mean number of responses for the clip. This typically involved between 5 and 20 keypress responses in a clip with only one hazard. The data from two participants had to be removed from accuracy and reaction time analyses, as they pressed the button too many times for most hazards, leaving insufficient data for analysis. This leaves an n of 24 in the unexposed group (12 in each speaking condition) and 20 in the exposed group (11 *exposed-silent* and 9 *exposed-block 1&2 speaking*) for accuracy and reaction time analyses.

A 2 x 2 x 3 ANOVA on the percentage of correctly identified predefined hazards (Figure 1, top row) revealed a significant main effect of block, $F(2, 80) = 3.59$, MSe = .123, $p < .05$, $p\eta^2 = .082$, with an overall increase in accuracy across consecutive blocks, and a significant interaction between training and speech, $F(1, 40) = 4.53$, MSe = .086, $p < .05$, $p\eta^2 = .102$. However, these were mediated by a significant three way interaction, $F(2, 80) = 4.20$, MSe = .123, $p < .05$, $p\eta^2 = .095$. To explore the source of the three way interaction simple interactions were calculated for each of the HP blocks. The interaction between training and speech was significant in block 1, $F(1, 120) = 14.11$, MSe = .111, $p < .001$, $p\eta^2 = .105$, and block 2, $F(1, 120) = 5.50$, MSe = .111, $p < .05$, $p\eta^2 = .011$, but not the final silent block, $F(1, 120) = 0.24$, MSe = .111, $p = .622$, $p\eta^2 = .002$. To explore the source of the simple interactions in the first two blocks, simple main effects were conducted for the effect of training at levels of speech in each block. In block 1, significantly more of the predefined hazards were correctly identified by exposed participants than unexposed in the always silent condition, $F(1, 120) = 9.07$, MSe = .111, $p < .01$, $p\eta^2 = .070$. However, exposed participants identified significantly *fewer* predefined hazards than unexposed participants in the speaking condition, $F(1, 120) = 5.38$, MSe = .111, $p < .05$, $p\eta^2 = .043$. In block 2 there was no significant effect of training in the always silent condition, but where participants were speaking, those who were exposed to commentary continued to identify significantly fewer hazards than unexposed participants, $F(1, 120) = 8.72$, MSe = .111, $p < .01$, $p\eta^2 = .068$.

These analyses address several research questions. The comparison of all four conditions in blocks 1 and 2 show commentary production after training reduces hazard perception accuracy, but commentary production without prior training has no effect on accuracy. The comparison of the exposed and unexposed silent groups across all three blocks shows that commentary training improves accuracy, but only in the first block after training. By block 3, in which we would expect to see the effects of training and practise on later silent hazard perception, there are no differences in accuracy between any of the groups.

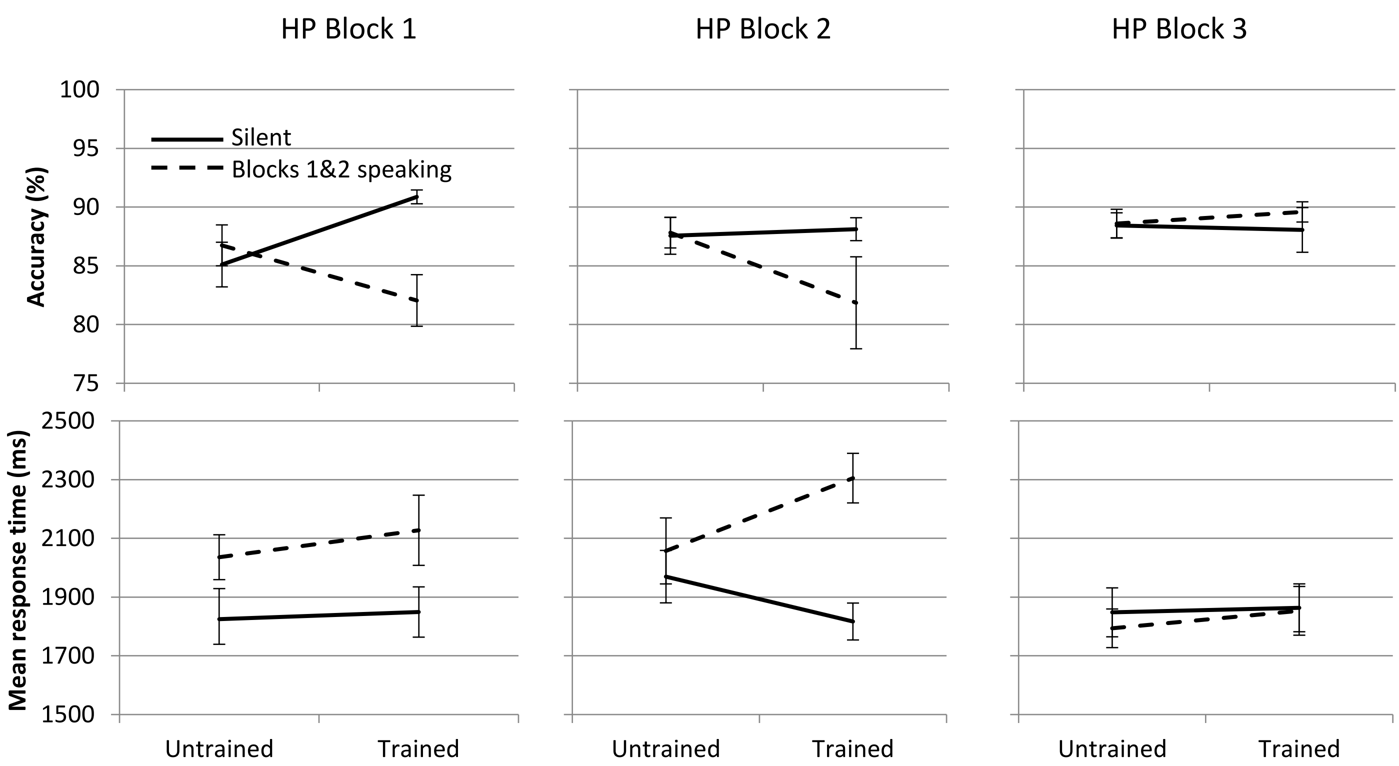

Figure 1. Percentage of predefined hazards that were responded to within the hazard window (accuracy %) and hazard response times (measured from the hazard onset) in HP blocks 1, 2 and 3. Dashed lines show those speaking in HP blocks 1&2, while solid lines show those participants who were silent throughout. All participants were silent in block 3. Error bars indicate one standard error above and below the mean.

3.2.2 Response times.

Response time results are shown in Figure 1 (bottom row). The 2 x 2 x 3 ANOVA yielded significant main effects of speech, $F(1, 40) = 5.09$, MSe = .897, $p < .05$, $p\eta^2 = .113$, and block, $F(2, 80) = 15.16$, MSe = 1.483, $p < .001$, $p\eta^2 = .275$, with generally quicker responses in the always silent conditions and in later blocks. However, these main effects were mediated by a significant interaction between block and speech, $F(2, 80) = 10.31$, MSe = 1.01, $p < .001$, $p\eta^2 = .205$, and a significant three way interaction, $F(2, 80) = 3.23$, MSe = .317, $p < .05$, $p\eta^2 = .075$. Simple interactions between training and speech at levels of block were calculated to explore the three-way interaction. In HP block 1 the interaction between training and speech was significant, $F(1, 180) = 7.09$, MSe = .124, $p < .01$, $p\eta^2 = .056$. Simple main effects of training at levels of speech showed that training increased response times in the blocks 1&2 speaking condition $F(1, 120) = 4.80$, MSe = .124, $p < .05$, $p\eta^2 = .038$, but did not affect response times in the silent condition. In HP block 2 the same pattern was observed, with a significant interaction between training and speech, $F(1, 180) = 11.08$, MSe = .124, $p < .05$, $p\eta^2 = .085$, caused by training increasing response times in the blocks 1&2 speaking condition, $F(1, 120) = 10.26$, MSe = .124, $p < .01$, $p\eta^2 = .079$, but not affecting response times in the silent condition. In HP block 3 there was no significant interaction.

If commentary training improves hazard perception response times, we would expect to see this in the silent conditions in blocks 1 and 2, or as a main effect in block 3. In the silent conditions, commentary training has no effect in any block. In block 3 there was no significant main effect of commentary training and no interactions, suggesting that commentary training did not improve response times, regardless of whether participants had practised their commentary in earlier blocks. Analyses of the speaking conditions demonstrated that producing commentary after exposure increases response times, but producing a naïve commentary does not.

3.3. Eye movement analyses

Four participants' eye movements could not be accurately tracked, so eye movement analyses are completed with an n of 11 each in the *unexposed-silent* and *exposed-block 1&2 speaking* conditions and an n of 10 each in the *unexposed-block 1&2 speaking* and *exposed-silent* conditions. Trials with missing eye position data for more than 10% of the clip were not included in the eye movement analysis (10% of 1276 trials).

3.3.1 General eye movements.

Mean fixation durations are shown in Figure 2. A 2 x 2 x 3 mixed design ANOVA on mean fixation duration revealed a significant main effect of block, $F(2, 76) = 10.68$, MSe = 1.218, $p < .001$, $p\eta^2 = .219$, with generally longer fixations in later blocks. However, this main effect was mediated by a significant interaction between block and speech, $F(2, 76) = 14.13$, MSe = 1.611, $p < .001$, $p\eta^2 = .271$. Simple main effects of speech at levels of block showed main effects of speech in block 1, $F(1, 114) = 12.82$, MSe = .251, $p < .001$, $p\eta^2 = .101$, and block 2, $F(1, 114) = 22.50$, MSe = .251, $p < .001$, $p\eta^2 = .165$, with shorter fixations for block 1 & 2 speaking than always silent, but no main effect of speech in block 3.

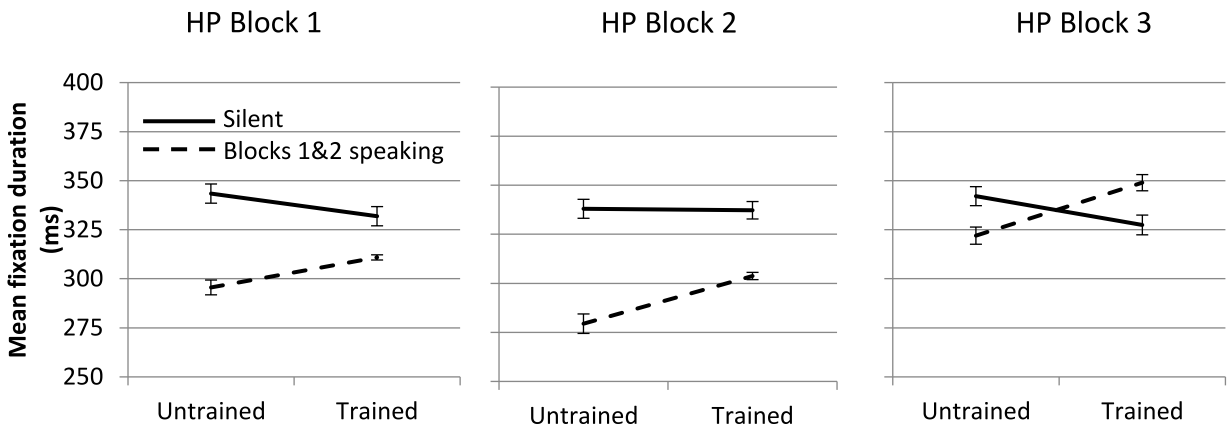

Figure 2. Mean fixation durations in the four between-subjects conditions in each of the three HP blocks. Dashed lines show those speaking in HP blocks 1&2, while solid lines show those participants who were silent throughout. Error bars indicate one standard error above and below the mean.

Analysis of fixation count data was consistent with that of fixation durations. Overall, these analyses show that, regardless of commentary training, commentary production causes shorter, more frequent, fixations.

Mean variance in horizontal and vertical fixation coordinates is shown in Table 2. A 2 x 2 x 3 ANOVA revealed no significant main effects or interactions between any factors for horizontal variance in fixation location. For vertical variance in fixation location there were significant main effects of training, $F(1, 38) = 8.68$, MSe = .921, $p < .01$, $p\eta^2 = .186$, and block, $F(2, 76) = 8.02$, MSe = .174, $p < .001$, $p\eta^2 = .174$. Those in the exposed group had significantly greater variance in fixation locations than those in the unexposed group. Difference contrasts revealed that the main effect of block was due to a significant increase in vertical variance in fixation location between the first and second blocks, $F(1, 38) = 16.49$, MSe = .282, $p < .001$, $p\eta^2 = .303$, but no difference between the first two blocks and the third block.

Table 2. Mean horizontal and vertical variance in fixation locations, time to first fixate hazard and first hazard fixation duration by condition. Time to first fixate hazard is expressed relative to the hazard onset time, but it was possible to fixate the hazard earlier in the clip. Negative values therefore indicate a fixation on the hazard causing object prior to the hazard onset. Standard errors are shown in parentheses.

	Always Silent		Block 1 & 2 Speech	
Block	Unexposed	Exposed	Unexposed	Exposed
	Horizontal Variance (Degrees visual angle squared)			
1	66.5 (7.0)	69.0 (6.4)	70.3 (8.9)	72.1 (4.7)
2	66.8 (8.6)	74.0 (5.2)	74.1 (8.8)	74.4 (6.5)
3	67.7 (9.9)	69.8 (14.6)	74.4 (8.0)	72.9 (8.7)
	Vertical Variance (Degrees visual angle squared)			
1	3.6 (0.7)	3.9 (0.6)	5.0 (1.4)	4.7 (1.0)
2	4.3 (1.1)	4.4 (1.2)	5.8 (2.2)	5.3 (1.1)
3	4.1 (0.9)	4.6 (1.5)	5.3 (1.4)	4.5 (0.6)
	Time to First Hazard Fixation (ms)			
1	1313.1	1187.4 (549.3)	-38 (435.9)	1200.1
2	1.2 (810.7)	-769.2 (460.2)	2322.3 (1220.5)	873.3 (521.1)
3	952.1 (586.4)	1006.8 (661.2)	-110.3 (704.8)	-830 (571.1)
	First Hazard Fixation Duration (ms)			
1	415.3 (26.5)	410.9 (31.7)	378.5 (23.2)	382.2 (13.8)
2	433.6 (21.1)	407.3 (18.9)	334.1 (22.4)	354.7 (20.6)
3	457.4 (36.4)	437.1 (28.3)	397.9 (11.3)	408.2 (16.8)

3.3.2 Fixations on the hazard.

An area of interest (AOI) was created to encompass the object that would eventually be at the centre of a hazardous event (vehicle/pedestrian) from the moment it entered the visual scene. Time to first fixate the hazard was calculated relative to the hazard onset (see Method) and was counted as the beginning of the first fixation in the AOI. A 2 x 2 x 3 ANOVA on time to first fixate the hazard (see Table 2) revealed a significant interaction between block and speech, $F(2, 76) = 8.97$, MSe = .116, $p < .001$, $p\eta^2 = .191$, with no other significant main effects or interactions. Simple main effects of speech at levels of block show no significant main effect of speech in block 1, but significant main effects of speech in block 2, $F(1, 114) = 11.48$, MSe = .105, $p < .001$, $p\eta^2 = .191$, and block 3, $F(1, 114) = 6.63$, MSe = 105, $p < .05$, $p\eta^2 = .073$. In the second block, block 1 & 2 speaking participants look at the hazard later than always silent participants. In the third block, the block 1 & 2 speaking participants look at the hazard sooner than the always silent participants.

Analysis of the mean duration of the first fixation on the hazard (first fixation within the AOI, see Table 2) showed a significant main effect of speech, $F(1, 38) = 8.26$, MSe = .112, $p < .01$, $p\eta^2 = .179$, with a shorter first fixation duration in the block 1 & 2 speaking condition than the always silent condition, and a significant main effect of block $F(2, 76) = 5.13$, MSe = .131, $p < .01$, $p\eta^2 = .119$. Planned difference contrasts showed that the effect of block was due to an increase in first fixation duration in block 3 relative to the first two blocks, $F(1, 38) = 8.00$, MSe = .224, $p < .01$, $p\eta^2 = .174$. These findings are in line with the analysis of eye movements across the clip, which showed shorter fixations overall when participants were speaking. 2 x 2 x 3 ANOVAs on number of fixations on the hazard and on total fixation time on the hazard showed no significant main effects or interactions.

Importantly, the analyses of eye movement data overall showed no indication of a beneficial effect of commentary training. The only effect of training on general eye movements was to increase vertical variance in fixation locations, with no effect on horizontal variance in fixation locations or fixation durations. AOI analyses showed no effect of training and no interaction involving training on any measure of eye movements relating to the hazard, meaning that there is no indication here that training affected hazard related eye movements.

4. Discussion

This study addresses a number of issues relating to the effects of commentary training and production on *concurrent* and *subsequent* hazard perception and eye movements. We compared the effects of producing a commentary after exposure with those of a naïve commentary and with silent hazard perception. We also investigated the effects of commentary *training* and *practice* on later silent hazard perception.

4.1. Subsequent effects of commentary training and practice

To examine the effects of commentary training on subsequent silent hazard perception, we can compare those exposed and unexposed participants who remained silent throughout all three blocks. There are no response time differences between these groups, but there is an improvement in accuracy in the first block of hazard perception clips in the exposed group relative to the unexposed group. With no benefit derived from commentary training during the second and third block of clips, it seems that any initial benefit of the current commentary training is either short-lived or obscured by a reasonably small amount of hazard perception practice for these drivers.

As this conclusion relies on the absence of a significant effect of commentary training, it is important to address the issue of statistical power to detect such an effect, if present. In a simple independent *t*-test comparing the exposed and unexposed group in the third block, which represents later silent hazard perception, there is no significant difference between the exposed and unexposed groups, $t(42) = .204$, $p = .839$, with a Cohen's d of 0.06.

To give this effect size some context, Cohen's d of 0.20, 0.50, and 0.80, are considered to represent small, medium and large effect sizes, respectively (Cohen, 1977) and to obtain Cohen's recommended power of .80 to detect an effect of this size would require 3,436 participants in each group (this was calculated using the power analysis program G*Power (Faul & Erdfelder, 1992; for a full description, see Erdfelder, Faul, & Buchner, 1996)). Other authors have reported effects of commentary that are very large. For example, Poulson et al. (2010) reported a Cohen's d of 1.54 when observing an effect of commentary training (n=10 in each group) and Isler et al. (2009) reported a post-training difference between a commentary exposed and control group with a Cohen's d of 1.04 (this was calculated based on the reported means and standard deviations, using G*Power). These very large effect sizes supported the conservative expectation of a large effect size of around Cohen's d of 0.8 in the current study. Our sample sizes led to a power to detect an effect of this size in an independent *t*-test of .83, which meets Cohen's suggested recommended power of .80.

In light of other work showing that commentary training improves performance in hazard perception tests, this is a surprising finding. Previous studies have typically involved an exposed group who watch a video filmed from the perspective of the driver, including a number of hazardous events, accompanied by an expert commentary. However, the control group typically watches the same video in silence (Horswill et al., 2010; Horswill et al., 2013; Isler et al., 2009; McKenna et al., 2006; Poulsen et al., 2010; Wallis & Horswill, 2007), alongside some additional non-hazard relevant safety information in some cases (Horswill, Falconer, Pachana, Wetton, & Hill, 2015; Wetton et al., 2013), but with no additional hazard perception task. While a silent control group eliminates the possibility that commentary benefits are a result of mere exposure to hazards, it does not create conditions in which both the commentary and control groups have *engaged* exposure to hazards. In the current study, both groups were asked to press a key when they saw a hazard in the training video, meaning that even those who were not exposed to a commentary were encouraged to engage in early hazard detection. The absence of an effect of commentary training relative to engaged hazard exposure suggests that commentary ensures trainee engagement with the road scene and developing hazards, but does not necessarily have hazard perception benefits beyond this. Meir, Borowsky and Oron-Gilad (2014) found that active training that involved increasing the trainee's experienced based knowledge, was superior to instructional training in improving hazard perception. In typical commentary training paradigms, the participants gain both active and instructional training (from watching commentary and searching for hazards, relative to a passive control), while the current study disentangles these two elements by allowing both groups to gain active training, from hazard perception practice, and adding only the instructional element in the commentary group. The finding that commentary does not add to the benefits of active engagement with hazards is in line with the finding that a hybrid of active and instructional training was not significantly better at improving hazard perception than active training alone (Meir et al., 2014).

Our findings do not necessarily undermine the potential utility of commentary driver training. All participants in the current study had held a UK driving licence for at least one year, and some had held a licence for much longer. Although some of the studies mentioned previously have shown improvements with very experienced drivers (licence holders for over 50 years on average: Horswill et al., 2010; Horswill et al., 2015; experienced police drivers: Horswill et al., 2015), the absence of separate experienced and novice driver groups in the current study means that it remains possible that learner or novice drivers may still benefit from video-based commentary training, relative to an engaged control group, but more experienced drivers do not. Future research should address this issue, with separate experience groups. This is particularly relevant, as Gregersen (1994) found that commentary practise may be overly demanding for very new drivers, so the unravelling of the effects of practise and production may be important for this vulnerable group. In addition, commentary training may be beneficial in the context of on-road training accompanied by an instructor. In this case, the content of a commentary may provide useful information to the instructor about the elements of the road scene that trainees are paying attention to, or more critically, the possible hazardous events that trainees fail to predict and prepare for. If watching a commentary training video engages trainees in the prediction of hazards sufficiently to improve their hazard perception skills, even if the same improvement could be obtained by actively searching for hazards, the driving public may find the commentary option more appealing than silent hazard searching, meaning that they are more motivated to engage in it. The negative effect of a live commentary on concurrent hazard perception is critical however, and gives a strong suggestion that trainees should not be encouraged to generate their own on-road commentary without supervision, though, it remains possible that trainees may compensate for the demands of a concurrent commentary while driving by reducing driving speed and prioritising driving responses over commentary production.

One explanation for the absence of any lasting beneficial effect of video-based commentary training would be that training alone is insufficient to obtain commentary benefits, and that it must be paired with commentary production, or even that commentary production in the absence of training is sufficient for benefits. However, in the current experiment, there are no differences in silent hazard perception speed or accuracy between those who had previously produced a commentary and those who had not. In this study, we asked participants to practise producing a commentary at the same time as providing hazard perception responses for analysis. One possibility is that the need to produce hazard perception responses while practising commentary production, undermined any potential learning from this practise. Future research could address whether there may be later benefits of practising commentary without the additional demands of making hazard perception responses.

Eye movement analyses allow us to investigate possible changes in visual search of the driving scene that commentary training and practice might cause, even if these changes do not lead to significant differences in behavioural measures of hazard perception. In particular, an increase in horizontal spread of search is an indication that the viewer is looking more widely for hazards and a broader horizontal spread of search has been observed in expert drivers than novices, for example (Mourant & Rockwell, 1972; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). In this case there was no effect of commentary on horizontal spread of search, though commentary exposed participants had a broader vertical spread of search. Broader vertical spread of search has been observed in novice drivers relative to those with more experience (Chapman & Underwood, 1998), so does not clearly indicate any improvement in visual search. It also cannot explain any detriment observed here, as detrimental effects are unique to exposed-*block 1 & 2 speaking* participants and increased vertical spread of search is common to exposed conditions, regardless of speech.

Commentary practitioners suggest that commentary training encourages trainees to look further down the road, or maintain a higher 'visual horizon' (Gilbert, 2007). Vertical spread of search may indicate that trainees respond to this by scanning up and down the road more, but behavioural results show that, in this case, this strategy is not successful.

4.2. Live commentary production

In Young et al. (2014) a live commentary was shown to impair concurrent hazard perception. This result was replicated in the current experiment; with exposed block 1 & 2 speaking participants taking approximately 400ms longer to respond to hazards in the blocks in which commentary was required. This gives a strong indication that producing a live commentary is detrimental for concurrent hazard perception. However, producing a commentary is a novel task, with which most participants would be unfamiliar. If the task is initially off-putting, but the distraction reduces with practice we would expect the detriment caused by a live commentary to be reduced in block 2 of the current study relative to block 1. However, this was not the case. There were no signs of practice having ameliorated the effects of a live commentary, suggesting that the difficulty caused by commentary is not limited to very early commentary generation.

In order to establish whether the effects observed when producing a live commentary would be caused by any speech about the road scene, or were specific to a commentary after exposure, we also compared the effects of live commentary production with and without prior commentary training. Increased errors and longer reaction times were only evident when participants were exposed to commentary before producing their own commentary. This suggests that there is something quite specific about the production of commentary after exposure that leads to slower responses to hazards. Concurrent verbalisation is similar to commentary driving in a number of ways, but, like the naïve commentary, it does not impair hazard perception (Crundall & Underwood, 1997). In concurrent verbalisation the driver is asked to say aloud what they are looking at and thinking while driving. This task is deliberately modelled on Ericsson and Simon's (1980) Think Aloud Protocol; designed to elicit information about participants' cognitive processes without interfering with a concurrent task. Given that concurrent verbalisation does not affect hazard perception, while commentary production after exposure does, the additional demands of commentary based on example, over and above the demands of concurrent verbalisation, are likely to be an important source of driver distraction. These additional demands are specifically, producing a continuous commentary, without pauses, and making predictions about the future status of elements of the environment (e.g. pedestrians/vehicles).

As the naïve commentary produced in the unexposed group did not lead to poorer performance, this may be due to them not meeting one of these two demands; either by limiting their commentary to observations, without predictions about the future behaviour of other drivers or by saying less/pausing more. However, reduced fixation durations associated with speech (discussed below), were encountered in both the exposed and unexposed block 1 & 2 speaking groups. If the *unexposed-block 1 & 2 speaking* group were producing enough speech to cause shorter fixation durations it seems likely that the amount of speech produced was comparable to those in the *exposed-block 1 & 2 speaking* group, suggesting that a lack of predictions in the unexposed group may explain the lack of a detrimental effect of speech. Future research should analyse recordings of commentaries, to allow direct comparison of the speech produced by the groups.

4.3. Eye movements

In Young et al. (2014) we showed that shorter fixation durations were observed when a live commentary was produced after training. However, these findings did not allow us to draw conclusions as to whether shorter fixation durations were caused by commentary training, commentary production, a combination of the two or a speaking task more generally. Experienced drivers show shorter fixation durations than novice drivers (Chapman & Underwood, 1998), so shorter fixations caused by commentary training, or even commentary practice, may give some hope that commentary is improving visual search behaviour. However, shorter fixations have also been observed in driving tasks, while carrying out non-spatial secondary tasks (Nunes & Recarte, 2002) or during naturalistic conversation (McCarley et al., 2004). Therefore, if shorter fixations are associated with all speaking conditions it is likely that they are caused by increased verbal workload, rather than a potentially beneficial effect of training. In the current study we found that block 1 & 2 speaking participants made more, shorter fixations than always silent participants, along with a consistent reduction in the length of the first fixation on the hazard. These differences were only present in blocks 1 and 2, where block 1 & 2 speaking participants were actually producing speech, with no difference in block 3. This was the case regardless of whether the participant was exposed to a commentary before the hazard perception test. This provides compelling evidence that shorter fixation durations are a result of speaking about the road scene at test, rather than a product of commentary training or the production of a commentary after exposure. These results are consistent with an increased verbal workload explanation, which does little to improve the outlook for commentary driver training.

There was no effect of commentary training or practice on number of hazard fixations or total time spent fixating hazards, suggesting that the detriment of commentary production cannot be explained by reduced attention to hazards and giving no indication of potentially positive changes that result from training in commentary that might lead to benefits too subtle for the HP test to measure. The first fixation on the hazard was later in the block 1 & 2 speaking group than the always silent group in the second block, and later in the always silent group than the block 1 & 2 speaking group in block 3 (where no speaking was required). This is a surprising finding, which cannot be easily explained. Combined with the finding that the length of the first fixation on the hazard is longer in the third block than previous blocks overall, this may suggest a complex fatigue effect with some differential effects depending on whether the participants have switched tasks or not between the second and third block.

4.4. Conclusions

This is the first time that the effect of commentary training on later hazard perception has been assessed relative to hazard perception practice. The results clearly show that there is no lasting benefit of video-based commentary training relative to engaged hazard exposure.

Results suggest that the key factor here is engaged observation of the road scene, whether this is through looking for hazards in silence or listening to a commentary.

In addition, the detrimental effect of commentary production has been reproduced, suggesting that drivers should not be encouraged to produce a commentary without expert supervision. This means that the potential dangers of commentary production should be made clear to new drivers who are asked to produce a commentary at test (New Zealand Transport Agency, 2016), to ensure that commentaries are never carried out without the proper supervision of an expert. When police advanced drivers are engaged in a high speed pursuit it is necessary to convey risk-related information back to the control room. While drivers are already advised to have the passenger do this when in a double crewed vehicle, the driver can provide the commentary in a single crewed vehicle (College of Policing, 2013). The effect of this practice under highly demanding conditions should be carefully investigated, given the detrimental effects of a commentary on hazard perception even without the additional demands of vehicle control.

Commentary training may still be useful as a method of ensuring that drivers actively engage in the search for hazards in order to improve later hazard perception, but there is no evidence to suggest that there are any additional effects of commentary exposure or practice per se, in the absence of an instructor.

5. 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

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.

Castro, C., Ventsislavova, P., Peña-Suarez, E., Gugliotta, A., Garcia-Fernandez, P., Eisman, E., & Crundall, D. (2016). Proactive listening to a training commentary improves hazard prediction. *Safety science, 82,* 144-154.

Chapman, P., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception, 27*(8), 951-964.

College of Policing (2013). *Road Policing: Police Pursuits.* Retrieved from: https://www.app.college.police.uk/app-content/road-policing-2/police-pursuits/

Coyne, P. (1997). *Roadcraft: The Essential Police Driver's Handbook.* London, UK: The Stationery Office.

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., Bains, M., Chapman, P., & Underwood, G. (2005). Regulating conversation during driving: a problem for mobile telephones? *Transportation Research Part F-Traffic Psychology and Behaviour, 8*(3), 197-211. doi:10.1016/j.trf.2005.01.003

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.

DIA. (2014). DIAmond Advanced Motorists: DIAmond Special Test. Retrieved from http://www.driving.org/diamond/tests/thetests

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. Behavior research methods, instruments, & computers, 28(1), 1-11.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological review, 87*(3), 215-251. doi:10.1037/0033-295X.87.3.215

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]. Bonn, FRG: Bonn University, Department of Psychology.

Gilbert, C. (Writer). (2007). Ultimate driving craft [DVD]. Available from http://www.driving4tomorrow.com/professional-driving-course.html.

Horswill, M. S., Falconer, E. K., Pachana, N. A., Wetton, M., & Hill, A. (2015). The longer-term effects of a brief hazard perception training intervention in older drivers. *Psychology and Aging, 30*(1), 62-67. doi:http://dx.doi.org/10.1037/a0038671

Horswill, M. S., Kemala, C. N., Wetton, M., Scialfa, C. T., & Pachana, N. A. (2010). Improving Older Drivers' Hazard Perception Ability. *Psychology and Aging, 25*(2), 464-469. doi:10.1037/a0017306

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.

Horswill, M. S., Taylor, K., Newnam, S., Wetton, M., & Hill, A. (2013). Even highly experienced drivers benefit from a brief hazard perception training intervention. *Accident; analysis and prevention, 52*, 100-110. doi:10.1016/j.aap.2012.12.014

Hughes, P. K., & Cole, B. L. (1986). Can the conspicuity of objects be predicted from laboratory experiments? *Ergonomics, 29*(9), 1097-1111.

IAM. (2014). Institute of Advanced Motorists: IAM Masters programme. Retrieved 4th March 2015 from http://www.iam.org.uk/drivers/motorists-courses/masters

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

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

Meir, A., Borowsky, A., & Oron-Gilad, T. (2014). Formation and evaluation of act and anticipate hazard perception training (AAHPT) intervention for young novice drivers. *Traffic Injury and Prevention, 15*(2), 172-180.

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

New Zealand Transport Agency (2016). *Full licence test guide (class 1).* Wellington: New Zealand Transport Agency. Retrieved from: http://www.nzta.govt.nz/assets/licence/photo/docs/full-test-guide.pdf

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

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. Psychological Bulletin, 116(2), 220-244. doi: 10.1037/0033-2909.116.2.220

Poulsen, A. A., Horswill, M. S., Wetton, M. A., Hill, A., & Lim, S. M. (2010). A brief office-based hazard perception intervention for drivers with ADHD symptoms. *Australian and New Zealand Journal of Psychiatry, 44*(6), 528-534.

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.

Spolander, K. (1990). Effects of commentary driving: A study on young male drivers. *VTI Rapport 359. Linköping: Swedish Road and Transport Research Institute*.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*(6), 629-646. doi:10.1080/0014013031000090116

Wallis, T. S. A., & Horswill, M. S. (2007). Using fuzzy signal detection theory to determine why experienced and trained drivers respond faster than novices in a hazard perception test. *Accident Analysis and Prevention, 39*(6), 1177-1185. doi:10.1016/j.aap.2007.03.003

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/cohortiifindings.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? *Accident Analysis & Prevention, 54*, 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. Hillsdale, NJ: Erlbaum.

Young, A. H., Chapman, P., & Crundall, D. (2014). Producing a commentary slows concurrent hazard perception responses. *Journal of Experimental Psychology: Applied, 20*(3), 285-294. doi:10.1037/xap0000016