Transportation Research Part F: Psychology and Behaviour 95 (2023) 432–449

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

Short term memory and peripheral vision at junctions

Peter Chapman *, Seda Orhan, Lily Moore

School of Psychology, University of Nottingham, Nottingham NG7 2RD, UK

A R T I C L E I N F O

Keywords:
Driving Behaviour
Gap Acceptance
Short Term Memory
Motorcycle
Peripheral Vision

A B S T R A C T

Motorcyclists are at extremely high risk of death in crashes where another vehicle pulls out into their path at a junction. Such crashes have often been described as the result of "look but fail to see" errors. However, recent research has shown that such errors occur even after a driver has looked directly at an oncoming motorcyclist. An alternative explanation for some of these crashes is the "saw but forgot" error. The idea that drivers might forget vehicles they have looked at only moments earlier is surprising, but it matches recent cognitive research that highlights limitations in visual short term memory. We present a cognitive model that highlights the limitations of drivers' vision and memory at junctions. In two laboratory experiments we explore 68 young drivers' abilities to make safe decisions at junctions and their memory for oncoming vehicles. In each experiment the driver is required to make head movements to view static images to the right, the left, and straight ahead. Views are projected over 180 degrees and the driver's basic task it to respond when they feel the junction is safe for them to cross. On occasional trials drivers are given a surprise memory test. Drivers find this task remarkably hard, frequently missing safe crossing opportunities and choosing to cross when it is not safe. Notably, drivers are much more likely to choose to a cross in front of an oncoming vehicle once they have looked away from it, suggesting that in this task failing to remember an oncoming vehicle is a more frequent cause of errors than failing to perceive it. This conclusion is supported by the memory tests which revealed that in even moderately busy traffic conditions the junction crossing task can exceed the capacity of visual short term memory. In the first experiment we separately explore the influence of peripheral vision, finding that drivers' decision making and memory in this task is poor with or without peripheral information being present. In the second experiment we explore an initial intervention designed to enhance phonological memory but do not find that it improves memory for oncoming motorcyclists. In the second experiment we additionally separate out the influence of load (number of vehicles present) and delay (how long ago the vehicles were seen). This separation shows clear effects of load, but not delay on memory. We interpret these as evidence for both proactive and retroactive interference in visual short term memory. The current research shows that even a simple junction crossing task can exceed the capacity of visual short term memory. We suggest that this may be an important cause of junction crashes.

1. Introduction

Motorcyclists are at particularly high risk of involvement in crashes on the roads. In 2021 310 motorcyclists died on British roads,

* Corresponding author.
E-mail address: peter.chapman@nottingham.ac.uk (P. Chapman).

https://doi.org/10.1016/j.trf.2023.05.004
Received 29 November 2022; Received in revised form 3 April 2023; Accepted 4 May 2023
Available online 26 May 2023

compared to 682 car occupants (Department for Transport, 2022), this ratio of approximately 2.2 car occupant fatalities to 1 motorcyclist fatality is remarkable considering that the ratio of miles travelled in England is currently approximately 190 by car to each 1 travelled by motorcycle (National Travel Survey 2022). A loose comparison of these numbers indicates that the risk of death per mile travelled as a motorcyclist might be in the order of 87 times higher than that as a car occupant. Recent changes in UK travel related to COVID-19 lockdowns do mean that data from recent years may be unusual, but the extremely high risk associated with riding a motorcycle has been a feature of UK and worldwide crash statistics for many years. Similar data from the United States (National Safety Council, 2022) reports that motorcyclists represented 18% of vehicle occupant deaths despite accounting for only 0.6% of the miles travelled in the country, while in the European Union 18.4% of road fatalities are motorcyclists or moped riders (European Commission 2020). Worldwide estimates suggest that motorcyclists (and other powered two and three-wheelers) may represent almost 30% of road deaths (World Health Organization, 2022). There are numerous reasons for high casualty rates for motorcyclists, including the motorcyclist's own risky behaviour and their physical vulnerability when a crash occurs, but the current research explores one specific type of common crash that seems to be the result of errors made by other road users. It has been estimated that more than 25% of fatal crashes in the UK involving motorcyclists happen when another road user moves into their path (Clarke et al., 2007) typically at a junction (Department for Transport, 2021). This is not just a finding observed in the UK but is seen internationally (Walton et al., 2013). Understanding why other road users might pull out in front of oncoming motorcyclists at junctions is thus a priority in preventing unnecessary deaths on the world's roads.

A common explanation for accidents involving motorcycles at junctions is that other road users (typically car drivers) fail to spot the oncoming motorcycle and hence pull out onto a junction into the path of the approaching motorcyclist who has no time to take any evasive action (Crundall et al., 2012). These have been labelled as 'Look but Fail to See' (LBFTS) crashes because it has been observed that car drivers may look in the direction of an oncoming motorcyclist but still decide to pull out in front of it (Brown, 2005). Crundall et al. (2012) investigated this by presenting participants with video clips of them approaching a T-junction across three screens and having them decide whether it was safe to pull out or not. Experienced drivers made shorter fixations on motorcycles than cars which may suggest that drivers did not notice that there was a motorcyclist approaching or they may have simply processed motorcyclists faster than cars. There have been alternative considerations of these LBFTS crashes. Robbins et al. (2018) investigated the gap that drivers leave before pulling out at a junction when a car or motorcycle was approaching and found that drivers left significantly smaller gaps in front of motorcycles than they did for cars. This is in line with previous suggestions that these accidents occur due to a size-arrival effect where smaller objects are considered to arrive later than larger objects suggesting that drivers may believe that motorcyclists are further away than they truly are, thus pull out when it is unsafe to do so (Horswill et al., 2005).

As will be obvious from the statistics cited above, motorcycles are much rarer than cars on UK roads and there is evidence that in countries where motorcycles are more common drivers are better at perceiving oncoming motorcyclists at distance (Lee et al., 2015). Such effects seem to be in addition to wider cross-cultural differences in visual search (Miller et al., 2021). The fact that motorcyclists are generally rare and thus likely to be unexpected has proved a key component of inattentional blindness (Mack, 2003) accounts of LBFTS crashes involving motorcyclists which suggest that because of their rarity motorcycles are afforded a lower level of attentional bandwidth than cars (Pammer et al., 2018).

These accounts of crashes at junctions have in common the idea that oncoming motorcyclists may never have been seen, or if they have been seen, perception or judgment of the motorcyclist has been in some way faulty. This has caused legislation to increase the visibility of motorcyclists, for example by the compulsory requirement for daytime running lights which has been estimated to have reduced motorcycle crash risk by 4–20% (e.g. Davoodi & Hossayni, 2015). However, it should be noted that the high levels of motorcyclist fatalities in the U.K. and Europe reported above come from after the widespread adoption of daytime running lights. The idea that motorcyclists should increase their visibility also seems to imply that they are at fault for not being seen. Motorcyclist advocacy groups often point out that it is the driver of the other vehicle who is at fault in these crashes and that interventions should be targeted at them. A problem with this approach is the mixed evidence on whether people can be reliably trained to improve their ability to spot oncoming motorcyclists (e.g. Crundall et al., 2017; Keyes et al., 2019; Olsen et al., 1981).

Although failures to see or attend to oncoming motorcyclists are certainly a major contributory factor in many of these junction crashes, a recent paper by Robbins et al. (2019) highlights the role of failures of visual short term memory (cf. Fagot & Pashler, 1995; Gunnell et al., 2019; Horowitz & Wolfe, 1998; Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997; Sperling, 1960; White & Caird, 2010). Robbins et al. (2019) found that when drivers pulled out into a junction in a high fidelity driving simulator they were often unable to remember key vehicles that were approaching them at the time. Critically, they found that drivers were particularly likely to fail to remember an oncoming motorcyclist, even when they had looked straight at the motorcyclist only seconds earlier. Of course a visual fixation on a relevant vehicle does not imply that it has been fully processed and there are a wide range of other measures that can be potentially used to assess attention in driving contexts (e.g. Lohani, Payne, & Strayer, 2019; Solis-Marcos & Kircher, 2019), however, Robbins et al. did find that the time since the last fixation (1.0 s versus 2.2 s) and the mean number of head movements made since that fixation (0.6 vs. 1.5 head movements) reliably predicted whether a vehicle would be remembered. This led them to suggest that "saw but forgot" errors for motorcyclists might be an additional cause of crashes at junctions.

A problem with investigating crashes at junctions is that although such crashes are alarming common at a statistical level, for individual road users they remain rare events, thus an average driver will pass through junctions perfectly safely thousands of times. This makes them difficult to investigate in the laboratory, indeed Robbins et al. (2019) found that actual crashes were extremely rare even with 135 drivers making multiple crossings of busy junctions. What they did find was evidence that drivers did not always have full information about what was around them at the time they started to cross the junction, however, even these memory errors were quite rare. Memory errors occurred on less than 10% of memory test trials, with many participants making no memory errors at all. Of course, even a rare memory error, leading to an even rarer crash, can remain absolutely critical to improving road safety – it doesn't matter that you can cross junctions safely a thousand times if on the next occasion your error causes a fatal crash. However this does make such errors hard to investigate experimentally, and it is a particular problem when testing involves the expensive and time-consuming use of an advanced driving simulator with simultaneous eye movement recording. Even though memory errors may be more common than the crashes that potentially result from them, investigating drivers' memory needs to be done with care. For results to have any potential application in the real world drivers' primary task must remain crossing a junction safely, rather than memorising oncoming vehicles. One of our main goals in the current research was to develop a paradigm where we could rapidly measure multiple crossing decisions at junctions and separately conduct carefully controlled tests of the drivers' memory at the time they made a crossing decision.

Robbins et al. (2019) proposed the PRC (Perceive, Retain, Choose) model of decision making as a cognitive model of the flow of information at the time a driver has to make a crossing decision at a junction. It contrasts with other frameworks for understanding junction crashes (e.g. Crundall et al., 2012) in that it highlights the dynamic flow of information over multiple head movements and the need to use visual short term memory in maintaining a representation of the environment. In Fig. 1 we present a revised version of this model that serves as the theoretical motivation for the present experiments.

The PRC model is based on the model of working memory that was initially proposed by Baddeley and Hitch (1974) and has been revised and expanded since then (e.g. Baddeley, 2010, 2012, Baddeley et al., 2012). The core idea of the model is that when a driver is at a junction they are engaged in central processing in order to make the decision to pull out (Go) or to wait (Accumulate and Update Evidence). The accumulation of evidence typically requires a head movement, which means that information that was previously directly available to vision is now only available from memory. Current versions of working memory (e.g. Baddeley et al., 2012) have incorporated a "hedonic detector" to account for emotional influences on attention and short term memory and the threat detection box in the PRC model was initially included to reflect this. However, in the original PRC model threats were assumed to be detected by drivers making a head and eye movement to focus directly upon them. In the revised version of the model shown in Fig. 1 we have explicitly noted the possibility that some threats might be detected in peripheral vision (pathway 4a). This incorporates the idea that we may be aware of a large or fast moving object even when we are not looking directly at it and raises the possibility that some information from peripheral vision may feed directly into a driver's decision making. A simple prediction from this aspect of the model would be that the relatively small size of motorcycles may make them less detectable in peripheral vision and thus less likely to be perceived as a threat when a driver is not looking directly at them. A critical feature of the revised model is that limitations of short term memory will mean that drivers sometimes forget vehicles they have previously looked at – when a large, high contrast, or fast moving object is forgotten it may still be detectable in peripheral vision, so although a memory error may occur it does not necessarily result in a crash because the threat can be peripherally detected. Even if an incorrect decision is made, there may be a later opportunity to abort the manoeuvre when a large vehicle becomes visible peripherally. Crashes involving motorcycles would thus occur not because they are more forgettable than other vehicles, but because when they are forgotten, the consequences are likely to be more severe.

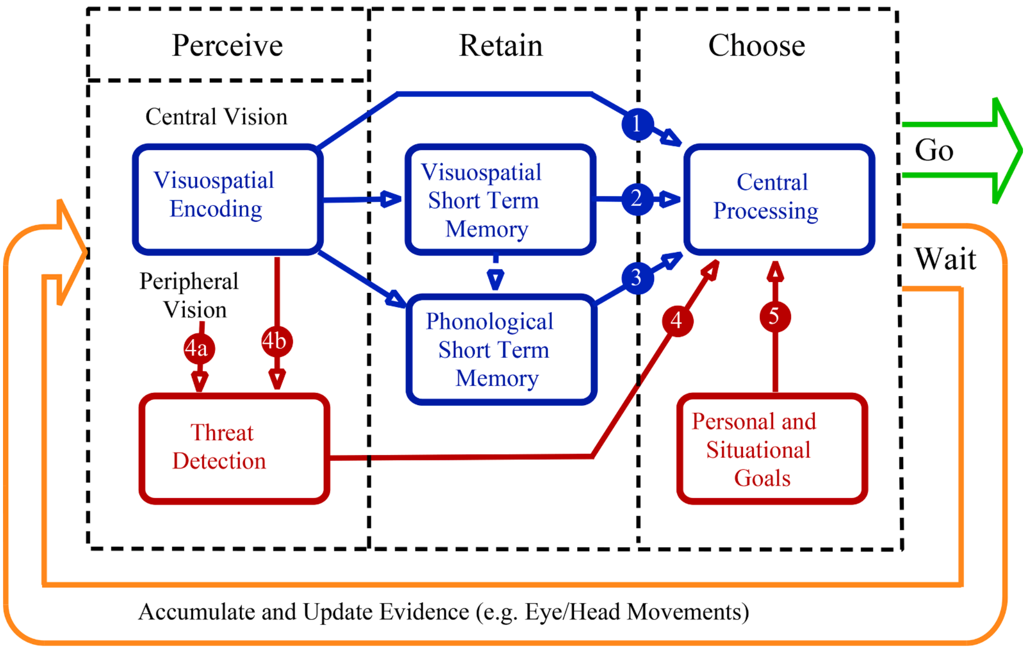

Fig. 1. A revised version of Perceive Retain Choose (PRC) model based on Robbins et al., (2019) with an additional pathway and an explicit comparison between central and peripheral vision. The model attempts to capture the fact that at the time when a decision is made a driver can only be looking in one direction and needs to combine perceptual information with information from short term memory. Each time a head movement is made, new perceptual information is available, but previous content needs to be stored if it is to be used in an updated decision.

The current research was motivated by a desire to explore the implications of the PRC model by designing a paradigm in which we could rapidly measure multiple decisions at a junction and conduct controlled tests of short term memory in a way that would not interfere with the drivers' decision making. Although our goal was to design a task that would not require participants to drive a high fidelity driving simulator, it remains critical that drivers are required to make multiple head movements and that the visual angles subtended by vehicles match those that would be found in real driving environments. We thus devised an applied version of a classic n-back task (Kirchner, 1958). In a classic n-back task a participant might listen to a sequence of letters and have to indicate if the current item was the same as one presented two items earlier (n-2). The reason such a task proves challenging is that each time a new item is presented the participant needs to fully update the content of working memory -encoding the new item into memory (n), comparing it with the relevant item (n-2), while still retaining the intermediate item (n-1) which will be relevant on the next trial, and presumably actively forgetting the now irrelevant earlier items from the list (n-3 and above). Our version of the task (like Kirchner's orginal one) is a visual one, but we believe that it is unique in that information has be retained and continuously integrated over the two previous head movements. The driver's primary task is to decide whether a junction is safe to cross based on the information they are currently

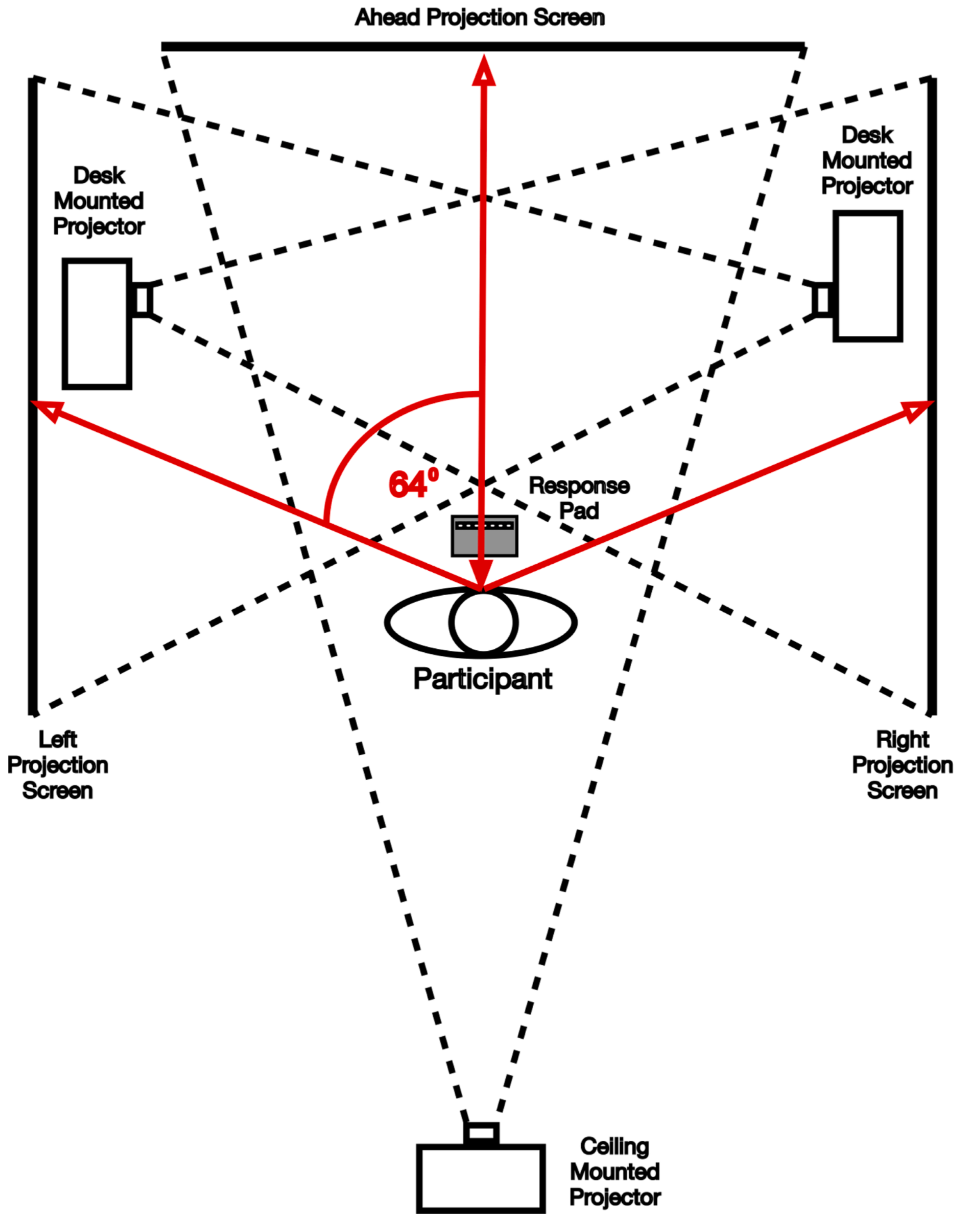

Fig. 2. The setup of the laboratory (drawn approximately to scale). Each projection screen is 2 m in width and the participant sits at a distance of 1.4 m from each of the three screens. The full viewing angle is thus in excess of 180 degrees but the middle of the left and right screens (where the centre of the road is displayed) is at a visual angle of 64 degrees from the view straight ahead.

looking at, combined with relevant information from the last two places they have looked. While this is not an exact analogue to the classic n-back task, it shares the need to integrate and update information in working memory and involves a process of combining information from multiple head movements that is critical to everyday driving.

Our prediction is that if memory failures account for errors at junctions then we would expect very few errors based on information on the screen a driver is looking at (current), but increasing errors when the relevant information comes from the previous head movement (n-1) or information from two head movements ago (n-2).

Although the main purpose of the task is to examine decision making, we can potentially obtain additional insights about the content and capacity of memory in this task asking drivers directly about the contents of any of the last three screens they have seen. Our expectation is that memory will be surprisingly poor, though we have no reason to predict that forgetting will be more dramatic for motorcycles than other vehicles. Where we might find differences between motorcycles and cars would be in their detectability, particularly in peripheral vision. We manipulate the availability of extreme peripheral vision by having two versions of the task – a full view version in which the whole visual world is continuously present, and a restricted view version in which only the direction in which the driver is looking is visible.

2. Experiment 1

2.1. Method

2.1.1. Design

The experiment has two phases, in the first phase we simply measure the decisions the driver makes about crossing the junction. In the second phase we additionally add some direct memory tests where we pause the task and probe the driver's memory for specific items. In each phase the basic design for the experiment is a mixed 2 × 2 design with the factors of Peripheral Vision (restricted view vs. full view) as a between participants factor, and Vehicle Type (car vs. motorcycle) as a within participants condition. For crossing decision errors in the first phase we also include the additional within participants factor of which Screen an oncoming vehicle was present on with three levels (current, n-1, n-2). For memory analyses from the second phase we instead include the within participants factor of Delay with two levels (previous screen vs. older screen) and of Direction (which direction a vehicle appeared to be travelling in) with two levels (towards vs. away).

2.1.2. Participants

These were 36 drivers, predominantly students at the University of Nottingham who were completing the experiment for course credit. There were 12 males and 24 females. Their mean age was 20.03 years. They all had a full or provisional UK driving licence and they had been driving for 2.18 years on average. Participants were assigned alternately to the full view and restricted view conditions, such that 18 participants took part in each condition. An a priori power analysis using G*Power3.1 (Faul et al., 2007) revealed that using 36 participants would give us a power of 0.91 to detect medium effect sizes (Cohen's $f = 0.25$) on within participants factors with three levels (such as Screen) and a power of 0.83 detect a medium effect size (Cohen's $f = 0.25$) on the interaction between the conditions of Peripheral Vision and Vehicle Type.

2.1.3. Apparatus

Static road scenes were created in the NITES (Nottingham Integrated Transport and Environment Simulation) facility and showed the view a driver would have when waiting to cross a suburban crossroads from a minor road, split into the view to the left, the view ahead, and the view to the right. The scenes were projected onto 3 large screens, see Fig. 2. A PC fitted with an AMD FirePro™ 2460 Multi-View quadruple head 512 MB graphics card running PsychPy3 version 3.1.5 (Peirce, 2007, 2009) was used to display images and record response taken on a Cedrus RB730 response pad. Images were displayed using three XGA projectors with each of the three images padded with a grey region above to leave the visible road image at a resolution of 1024 pixels horizontally and 543 pixels vertically to match the true aspect ratio of the simulated scene. Because we used three flat screens without warping or blending there was a small area near to where the interior pillars of a car would be where no image was projected and the image near to these areas was distorted, however, resolutions and distances were carefully chosen such that vehicles projected on the roads to the left, ahead, and right were undistorted and subtended precisely the same visual angles that they would in the full NITES simulator and indeed in the real world, thus an oncoming car subtended a visual angle of approximately 10° horizontally, while an oncoming motorcycle would subtend a visual angle of approximately 3° horizontally. The heights of cars and motorcycles were similar – subtending approximately 5° vertically.

2.1.4. Procedure

Participants viewed the road scene continuously, making repeated head movements to view the screen to their left, then the one straight ahead of them, then the one to their right, then back to the left screen again, and so on continuously. Each time they responded to one screen, the content on the screen they were about to look at updated. In the full view condition the content on the screen they had just responded to and that on the previous screen remained unchanged. In the restricted view condition the content of both of these screens was replaced by a version which had been blurred such that no vehicle information was visible. Each screen they looked at represented one trial, and they always had to make a decision (cross/wait) before moving onto the next screen. The whole experiment involved them viewing a total of 645 screens, starting with a screen to their left, and ending 644 trials later with a screen to their right, having viewed each direction 215 times in total. The first 45 trials (15 in each direction) served as practice block while the participant got used to the task. These were followed by a block of 432 trials (144 in each direction) which formed the main experimental block. After this there was a final block of 168 trials during which the participant continued to perform the cross/wait task but surprise memory tests were inserted on 12 occasions.

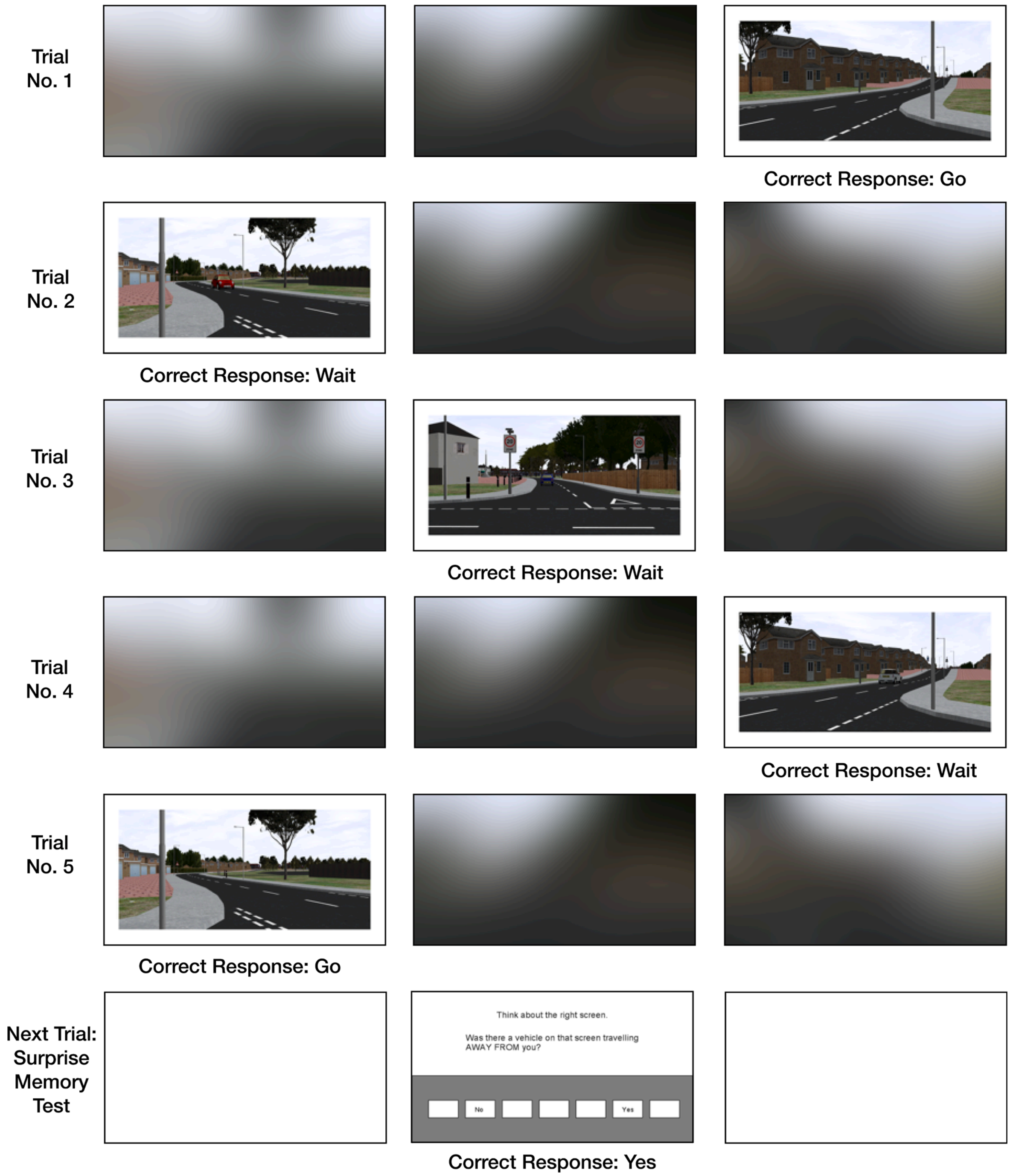

Fig. 3. A sample sequence of five trials from the experiment in the restricted view condition followed by the start of a typical memory test. In trial 1 of this sequence the participant is looking at the right hand screen and because there are no oncoming vehicles on this screen or the previous two the correct response is to go (assuming no oncoming vehicles had been present on the last two screens). On trial 2 an oncoming vehicle appears on the screen they are now attending to (the left screen) so the correct response becomes to wait. This remains the correct response until trial 5 when the participant is again looking at the left screen but it has become safe to go. Immediately after trial no.5 in this sequence the participant receives a surprise (n-2) memory test. The second illustration shows the same trial sequence as experimenced by a participant in the full view condition.

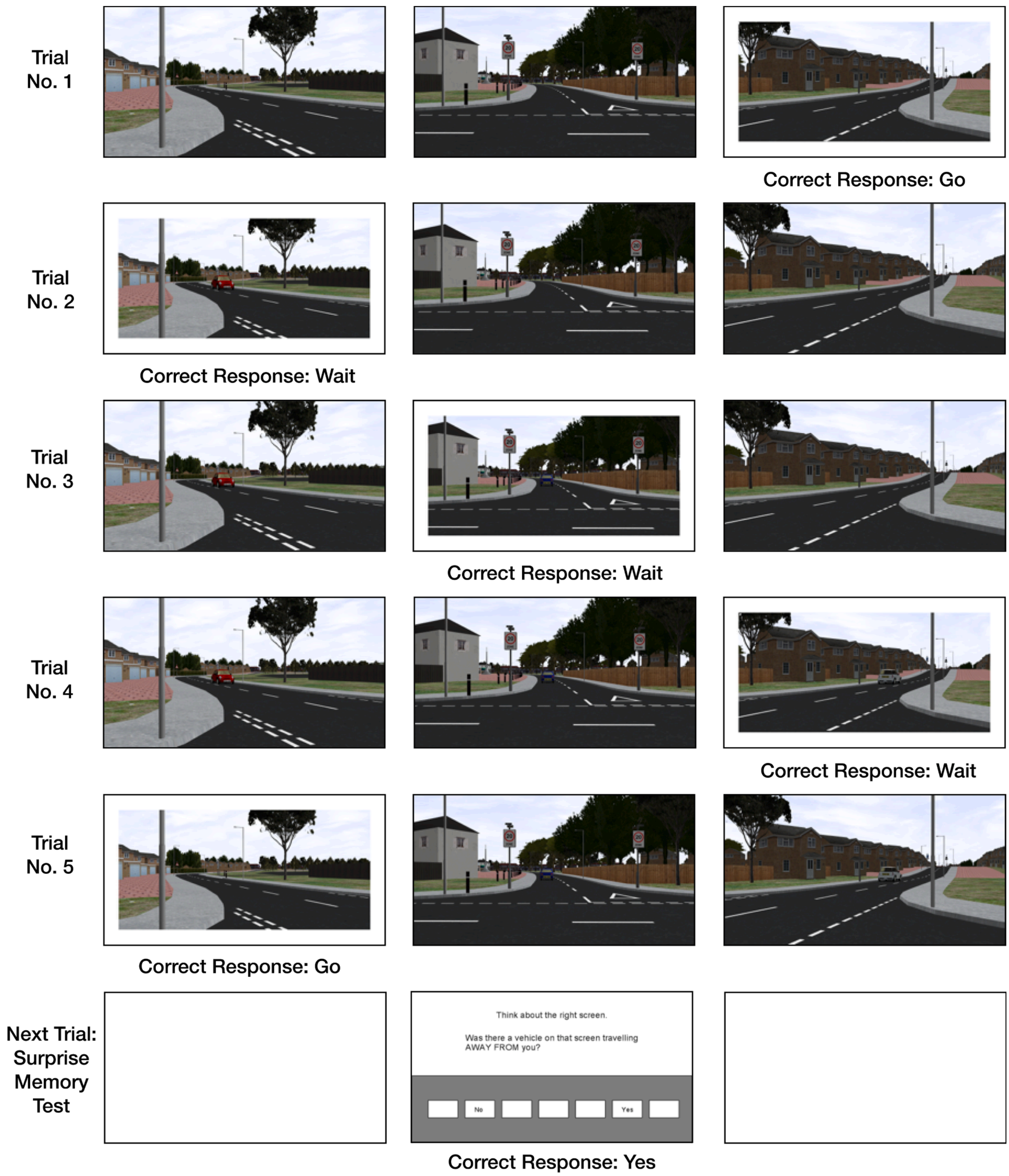

Fig. 3. (*continued*).

2.1.4.1. Wait/Cross decisions. The drivers' primary task throughout was to decide whether it was safe to cross the junction on each occasion based on the following rule: "It is only safe to cross if the screen you are looking at, and the last two screens you have looked at contained no oncoming vehicles (i.e. there is no oncoming traffic in any direction)." Participants were told that vehicles travelling away from them did not make it unsafe to cross, but that any oncoming traffic (including vehicles from the road ahead) meant that it was not safe to cross. For drivers in the full view condition the screen they were meant to look at on each trial was indicated by a white border around the road scene. For those in the restricted view condition the same border was always present on the screen to be viewed but the other two screens were replaced by a luminance equivalent blur. The level of blur was chosen to give the subjective impression that peripheral vision might be present, but in fact the blurred images were all identical and contained no vehicle information. Fig. 3a thus represents a series of five trials that a participant might see in the restricted view condition followed by a memory test. In the full view condition (Fig. 3b) the previous two screens remain present even while the driver looks at the screen with the white border. Although the examples in Fig. 3 includes a memory test, during the actual experiment no memory tests would be conducted until at least 477 cross/wait decisions had already been made.

Although there are always three screens displayed, the visual angles in the laboratory exactly match those at real junctions in Nottingham (Robbins, Allen & Chapman, 2019) and mean that the participant has to make combined head and eye movements of approximately 64° between the left and ahead screen, and 64° between the ahead and right screen, followed by a overall movement of approximately 128° to return from the right screen to the left screen. Although we did not separately measure head and eye movements in the current study we suspect that they matched the Robbins, Allen & Chapman (2019) study in which movements to the left and right were typically achieved with approximately 2/3 of the visual angle being covered by the head movement, with the remaining third being covered by an eye movement. The screen they should be looking at and responding to is always marked by a small white border, though in practice this is not really necessary after the first few trials since the participant develops a simple routine of looking left, ahead, right, left and so on, making a new head movement immediately after each response.

Images were always displayed for a minimum of 2 s and responses were only recorded after an image has been on the screen for at least 500 ms. If the participant had made a response between 500 ms and 2 s the trial ended after 2 s and the next trial began immediately, otherwise the screens remained present and unchanged until a response was made. This meant that participants obtained no particular benefit from trying to make speeded responses and the main experimental cross/wait block always took at least 14 min 24 s to complete. This timing was adopted after extensive piloting during which we discovered that alternative approaches could lead to missed trials or ambiguities about which screen the participant was responding to.

All possible combinations of vehicle type, colour, and combination were created and displayed in a partially randomised order that was balanced such that there were exactly 54 occasions amongst the 432 experimental trials where the correct response would be cross. This exactly matches the expected number, given that the probability of an oncoming vehicle was set at 0.5 on any trial, and that three successive screens without an oncoming vehicle were necessary for it to be safe to cross. For the remaining 378 trials there was an oncoming vehicle on at least one of the current three screens so the correct response would be to wait.

2.1.4.2. Memory tests. Once participants had completed the 432 experimental trials the experiment continued seamlessly into a memory block. Participants' primary task remained to determine whether it was safe to cross on each trial, but on 12 occasions in this block the experiment was interrupted by a surprise memory tests. The memory instructions always appeared instead of the next screen that would normally have been viewed and identified one of the previous screens. Half the time (6 tests) drivers were first asked whether there was a vehicle on a specific screen travelling towards them, while for the other six tests they were initially asked whether there was a vehicle on that screen travelling away from them. Drivers thus answered two questions about the *presence* of vehicles (towards and away) in a counterbalanced order. Because they were always tested for vehicles in both directions we can think of this as 24 memory tests (12 individual tests, with two defined locations in each test. For 12 of these 24 tests there actually was a vehicle present. Participants used the Cedrus response pad to indicate Yes/No regarding the presence of each vehicle. This information about presence is the potentially safety–critical information so was always probed first. Where they had responded "Yes" to the presence of a vehicle, after the two questions about presence they then had an additional question to answer about the vehicle type (car vs. motorbike) and that vehicle's colour (Blue, Red, or White). Responses were strictly 2AFC for presence, then 2AFC for type, and finally 3AFC for colour. The true frequency of vehicles were balanced such that across the whole experiment vehicles were present at exactly 50% of locations, and when a vehicle was present it had a 50% chance of being a motorbike, and each vehicle appeared equally often with each of the three colours (e.g. 33% of cars were Blue, as were 33% of motorbikes).

2.2. Results

Where analyses are conducted on factors with three levels of a repeated measure factor (i.e. screen), Mauchly's test of sphericity was conducted. On occasions where sphericity assumptions were breached, we have used a Greenhouse-Geisser correction, this is apparent in the reported degrees of freedom which will be decimals in these cases. Where we are comparing performance relating to the current screen with the two previous ones (cross/wait task) we have used a priori Helmert contrasts to first compare the current screen with the average of the two previous ones, and then to compare screen n-1 with screen n − 2.

2.2.1. Wait/Cross decisions

On average drivers chose to wait on 4.0 (range 0 to 31) occasions when it was in fact safe (missed opportunities) and chose to cross 25.8 times (range 0 to 114) when there was an oncoming vehicle on one of the three screens (dangerous crossings). Overall this performance equates to an average of 93.1% correct decisions (range 73.1 to 100). Generally people made similar numbers of correct decisions in the full view (mean = 92.4%) condition and the restricted view condition (mean = 93.8), [$t(34) = 0.574$, $d = 0.191$, $p = .57$]. There was no difference between the frequency of missed opportunities (0.075) and dangerous crossings (0.068) when each was expressed as a proportion of the available opportunities [$F(1,34) = 0.067$, $\varepsilon^2_{(p)} = 0.002$, $p = .797$], and the balance of errors types did not change as a function of Peripheral Vision [$F(1,34) = 0.468$, $\varepsilon^2_{(p)} = 0.014$, $p = .499$].

Participants generally found that the 2 s presentation for each screen was sufficient for them to make a response, and typically they did not make head movements away from the screen they were attending to until they had pressed the response key. Out of the 432 trials, on average 351 were completed within the 2 s window, and this did not differ between and the restricted view (mean = 357) condition and the full view (mean = 345) condition [$t(34) = 0.536$, $d = 0.179$, $p = 0.596$]. Where participants did go beyond the minimum 2 s viewing time the average extra viewing time was 1.02 s and this did not differ between conditions [$t(32) = 1.057$, $d = 0.363$, $p = 0.298$].

For most of the dangerous crossings there was just a single oncoming vehicle on one of the current three screens. For these cases we can calculate a dangerous crossing rate by dividing the number of dangerous crossings by the number of times the driver viewed such a combination. This proportion was analysed in a 2x2x3 mixed ANOVA using within participants factors of Vehicle Type (car vs. motorcycle) and Screen (current, n-1 and n-2) and the between participants factor of Peripheral Vision, see Fig. 4. The main effect of Screen was significant $F(1.11,37.60) = 11.62$, $\varepsilon^2_{(p)} = 0.255$, $p = .001$. Helmert contrasts revealed that dangerous crossings in front of an oncoming vehicle on the current screen (mean proportion = 0.063) were rarer than those to screens held in memory, $F(1,34) = 11.77$, $\varepsilon^2_{(p)} = 0.257$, $p = .002$, and that dangerous crossings were more common in front of vehicles viewed two screens previously (mean proportion = 0.218) than one screen ago (mean proportion = 0.106). There was no main effect of Peripheral Vision [$F(1,34) = 0.26$, $\varepsilon^2_{(p)} = 0.008$, $p = .613$] or Vehicle Type [$F(1,34) = 1.677$, $\varepsilon^2_{(p)} = 0.047$, $p = .210$] and no higher level interactions were significant.

2.2.2. Memory dependent Variables:

Our initial analysis focuses on the first question people were asked ("Was there a vehicle present?") which we explore as a function of Peripheral Vision, Delay, and Direction. There are two types of error possible in memory – here a 'miss' would mean forgetting to report a vehicle that was there, while a 'false alarm' would be claiming to remember a vehicle when none was in fact present. We have included Error Type as an additional within participants factor with two levels (miss vs. false alarm). There was a main effect of Error Type $F(1,34) = 16.40$, $\varepsilon^2_{(p)} = 0.325$, $p < .001$, with false alarms (mean proportion = 0.398), generally being more common than misses (mean proportion = 0.227). There were also significant interactions between Error Type and both Delay, $F(1,34) = 11.88$, $\varepsilon^2_{(p)} = 0.259$, $p = .002$, Direction, $F(1,34) = 21.67$, $\varepsilon^2_{(p)} = 0.389$, $p < .001$, and Peripheral Vision, $F(1,34) = 9.35$, $\varepsilon^2_{(p)} = .216$, $p = .004$ (see Fig. 5).

To provide a formal analysis of memory for vehicle type that combines hits and misses we calculate the proportion of a particular response type that was incorrect for each participant (having excluded 3 participants who did not use all the response options). A 2x2x2 Analysis of Variance on the two within participants factors of Vehicle Type and Direction plus the between subjects factor of Peripheral Vision yielded a significant interaction between Vehicle Type and Direction, $F(1,32) = 7.325$, $\varepsilon^2_{(p)} = 0.186$, $p = .011$ see Fig. 5(d).

A final question that we can ask of the memory data is how likely people are to actually remember full details of a vehicle that they have viewed. For this analysis we report the proportion of occasions on which people were able to accurately remember both the type and colour of a vehicle that was present on one of the screens. Overall participants were able to do this on 47% of occasions when the vehicle was travelling towards them, and just 25% of occasions when it was travelling away. A 2x2 Analysis of Variance with Direction and Peripheral Vision reveals only a significant main effect of Direction, $F(1,34) = 32.052$, $\varepsilon^2_{(p)} = 0.186$, $p < .001$.

2.3. Discussion of experiment 1

The experiment successfully created a challenging task, where drivers were generally able to make correct decisions but where most drivers made at least some errors. Some of these errors came in the guise of missed opportunities to cross the junction (approximately 7% of opportunities to cross were missed). The other type of error came in the form of dangerous crossing decisions – where the driver said that it was safe to cross while there was still an oncoming vehicle approaching. These errors were also made on

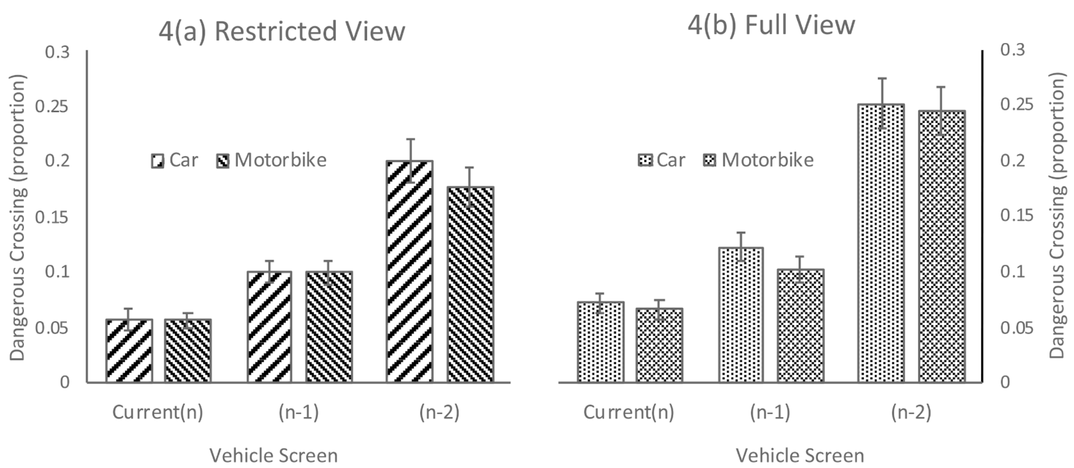

Fig. 4. The mean proportion of trials with a single oncoming vehicle on one of the three relevant screens where the driver incorrectly said that it was safe to cross (dangerous crossings). Data are split by what type of vehicle it was (car vs. motorbike) and whether the oncoming vehicle was on the screen the driver was currently looking at (n) or the one they looked at last (n-1) or the screen before that (n-2). Panel (a) shows data from participants where peripheral information was removed from screens n-1 and n-2. Panel (b) shows data from participants where full peripheral information remained present. Error bars show one standard error above and below each mean.

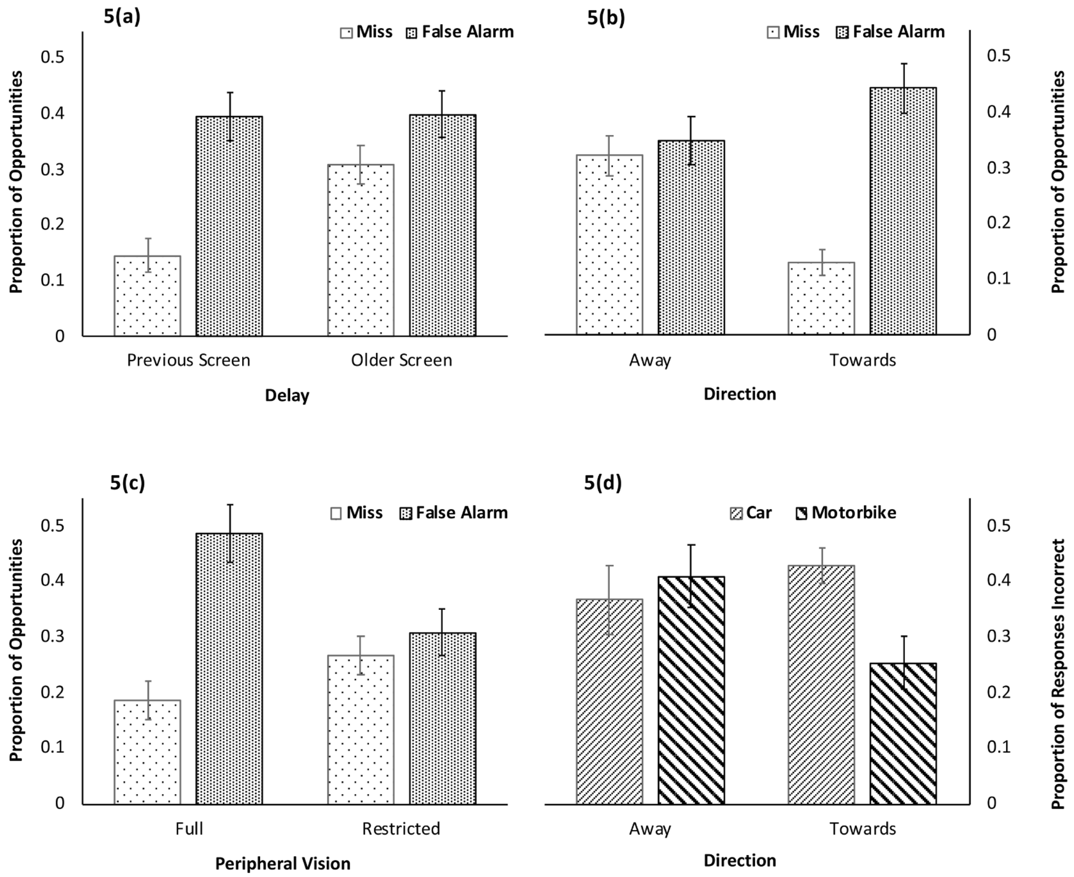

Fig. 5. Memory errors: The proportion of vehicles forgotten (misses) and incorrectly recalled (false alarms) as a function of Delay (panel a), Direction (panel b) and Peripheral Vision (panel c). Panel (d) shows the proportion of responses of a specific Vehicle Type (car vs. motorbike) that were incorrect as a function of Direction of travel. Error bars show one standard error above and below the mean.

approximately 7% of trials, and were much more likely to occur when the oncoming vehicle was on a screen that the driver was not currently looking at. Where a single oncoming vehicle was present two screens ago this false alarm rate rose to 22% of relevant occasions. For anyone doing this task these results would not be surprising – it feels hard to integrate information over three trials – but it is worth realising that the results in Fig. 4 seem imply that many of the vehicles a driver has successfully waited for when they looked at them have been forgotten or ignored a few seconds later when they are viewing the following screen, and even more have been forgotten or ignored if a second head movement has been made between viewing and responding. If we treat this a forgetting, we could estimate that while 93.7% of oncoming vehicles are correctly perceived, only 89.4% are remembered after a single had movement, and only 78.2% after a second one. This result is broadly comparable to the memory results in which participants correctly reported that an oncoming vehicle was present 94.4% of the time after a single head movement, and 79.2% of the time after two or three head movements.

It should be noted that this high level of errors comes about because we have created a task where memory about relevant vehicles from two head movements ago is repeatedly important and memory performance has to be mainained over 432 continuous trials. Although this level of short term memory load might be occasionally necessary in real driving, we would not expect errors in real driving to be anything like this frequent because most of the things a driver looks at will not affect their decision to cross a junction. The memory results from experiment 1 also show just how little information participants retain if it is not relevant to the current task (cf. inattentional blindness - Simons & Chabris, 1999). Memory for vehicles travelling away from the junction was notably poor (76.3% correct after just one head movement) and memory for vehicle details was even worse - overall we found only a 25% chance that a participant would be able to correctly report the type and colour of a vehicle travelling away from them.

The Peripheral Vision manipulation has very little effect on performance in either the cross/wait task, or the memory task. The only influence of peripheral information seemed to be to increase the number of false alarms made in memory, perhaps because participants felt that this condition should have helped their memory for vehicles and thus biased them to guess that vehicles were present more often even though it did not seem to actually improve their memory. Our original hypothesis from the revised PRC model (see Fig. 1) was that cars might be more visible in peripheral vision than motorcycles which might explain why crashes are more frequent with motorcycles and potentially why memory for motorcycles is more important than memory for cars. The results from experiment 1 provided no evidence to support the idea that peripheral vision is being used in these situations and it may well be that both cars and motorcycles in the current experiment were completely undetectable at 64° of peripheral angle (cf. Simpson, 2017; Wolfe et al., 2017). Note that although there is a very limited availability of information in peripheral vision for static stimuli, evidence suggests that for moving stimuli at around 70° of visual angle we may have relatively good sensitivity (To, Regan, Wood & Mollon, 2011). It thus becomes of considerable interest to attempt to create a version of this experiment where dynamic videos rather than still images are used.

The results from experiment 1 seem to clearly demonstrate that drivers' memory for vehicles they have previously looked at is surprisingly poor and that peripheral vision is not sufficient to prevent them pulling out in front of vehicles they are no longer looking at. Two aspects of the results from experiment 1 deserve further exploration. Although we were able to demonstrate frequent dangerous crossings that might relate to junction crashes in the real world we did not find any major differences in either crossing behaviour, or memory between cars and motorcycles. On balance we found memory for motorcycles to be potentially better than that for cars, at least when they are travelling towards the junction (see Fig. 5d). Of course, many authors have speculated that it is the infrequency with which motorcycles are seen relative to cars that accounts for people's failure to see or remember them (e.g. Pammer et al., 2018). In experiment 1 we have deliberately used equal frequencies of cars and motorcycles to ensure that any differences could not be attributed to frequency effects. In experiment 2 we reduced the frequency of motorcycles to a more representative figure of 1 motorcycle in every 20 vehicles. We chose to also include pedal cycles at the same frequency, with the expectation that explanations based on static vision and memory would apply equally to pedal cycles and motorcycles, and given the evidence that pedal cyclists are also the victims of LBFTS crashes (Herslund & Jørgensen, 2003). Nonetheless, the primary motivation for the inclusion of pedal cycles is simply to keep the frequency of motorbikes low, while still having enough trials for the analysis. The fact that we now have infrequent "bikes" (one every 10 screens on average) also makes it possible to conduct a preliminary exploration of the "see bike, say bike" intervention proposed in Robbins et al. (2019). They suggested that saying "bike" whenever a motorcycle is seen at a junction should automatically encode it into phonological short term memory – effectively increasing the capacity of short term memory. In experiment 2 we explore the possibility that this kind of intervention might affect both crossing behaviour and short term memory.

Experiment 1 did seem to show that memory was worse after additional head movements, but did not provide much detail of the nature of this effect. It may seem obvious that memory would decay over time, but this statement turns out to be surprisingly controversial. In fact memory theorists have often suggested that there is no such thing as decay in memory (e.g. Surprenant & Neath, 2009; Ricker et al., 2016), and that apparent decay is in fact simply the result of mixtures of proactive and retroactive interference (e.g. Keppel & Underwood, 1962). In our memory task successful performance would require people to hold all the information from three screens in visual short term memory simultaneously, so it is possible that poor performance is because of interference from the other items in short term memory and this could be both proactive and retroactive (e.g. when trying to remember an item on screen n-2, proactive interference from screen n-3, and retroactive interference from screen n-1 could both influence performance). If retroactive interference was dominant this might produce a delay effect (poorer performance for older screens), but if proactive interference is just as important then it may not matter whether interfering items come before or after the target item. We would expect poor performance when there were many other vehicles to remember, and good performance when there were few, but it would not matter which of the three relevant screens we chose to test. To explore this idea directly in Experiment 2 we chose to use memory tests that were exactly balanced in both delay and total memory load to see whether decay or interference provided a potential explanation for the levels of forgetting that were observed.

3. Experiment 2

To give people time to get used to the main cross/wait task before adding a secondary task and to improve our power to detect any effects of the see bike say bike task we chose to split the cross/wait part of the experiment into two matched blocks with the first (baseline) block being the same for all participants. We added a short break in the middle of the cross/wait task after which participants in the see bike say bike condition were asked to verbally say the word "bike" every time they saw either a motorcycle or pedal cycle, irrespective of the direction it was travelling. Participants in the see bike say bike condition then continued with this secondary task throughout both the second (experimental) block of cross/wait trials, and the block of memory trials. The main changes from experiment 1 to experiment 2 are thus the reduction in the frequency of bikes, the removal of the restricted view condition and the addition of the "see bike, say bike" secondary task for half the participants. A more subtle change is the exact balancing of load in the memory tests so we can systematically compare high and low load trials.

3.1. Method

3.1.1. Design

Experiment 2 is split into three phases, a first baseline block of cross/wait trials, a second experimental block of cross/wait trials, and a third block of cross/wait trials in which participants have occasional memory tests. The basic design for the first block is 3x2 within participants design with factors of screen (now, n-1, and n-2) and oncoming vehicle type (car vs. bike – because of the low frequency of bike trials, here we combine data from both pedal cycles and motorcycles). The basic dependent variable for this and the second block is the proportion of trials on which people incorrectly say that it is safe to cross (false alarms). For phase 2 of the experiment we expand this to a 3x2x2 mixed design with the additional between participants factor of secondary task (say bike vs. none). This can also be treated as a 3x2x2x2 mixed design by additionally comparing the results with those from phase 1 in an additional within-participants factor of block (baseline vs. experimental).

The phase 3 memory tests can be analysed two different ways. If we simply score whether drivers' responses to the question "was there a vehicle present?" was correct and combine data across vehicle absent and vehicle present trials we have a mixed 2x3x2 design with memory load (low vs. high) and delay (n-1, n-2, n-3) as within participants factors, and Secondary Task as a between participants factor (say bike vs. none). Alternatively, if we limit the analysis to occasions when a vehicle was actually present, the design is a 3x2x2 mixed design with vehicle type (car, motorcycle, or pedal cycle) and direction (travelling towards the junction, or away from the junction) as within participants factors and secondary task (say bike vs. none) as a between participants factor.

3.1.2. Participants

These were 32 drivers, predominantly students at the University of Nottingham who were completing the experiment for course credit and had not taken part in the previous experiment. There were 8 males, and 24 females. Their mean age was 21.2 years. They all had a full or provisional UK driving licence and they had been driving for 3.67 years on average. Participants were assigned alternately to the control and say bike conditions, such that 16 participants took part in each condition.

3.1.3. Apparatus/Stimuli

Road scenes were identical to those used in Experiment 1 except that we used NITES1 to additionally create images of pedal cycles in the same locations and combinations as cars and motorcycles in the previous experiment. Because the NITES vehicle models did allow pedal cycles to be clearly differentiated in terms of colour all pedal cyclists were shown wearing a grey t-shirt. For the memory test we modified the response options to allow participants the option of choosing car, pedal cycle, or motorcycle on the CEDRUS response box, and removed the requirement to identify vehicle colour. All other aspects of the apparatus were identical to Experiment 1, except for the precise order of trials presented and the timing of the memory tests. For experiment 2 we additionally counterbalanced and explored the factor of load – here defined by the number of additional irrelevant vehicles that were present at the time a decision was made.

Whereas in the previous experiment we had the frequency of cars and motorcycles equal, in experiment 2 the frequency of bikes was chosen such that bikes were rare compared to cars, thus out of every 20 vehicles 18 would be cars, 1 would be a motorcycle and 1 would be a pedal cycle. Since we kept the frequency of vehicles overall at 50% this meant that on average people would need to complete 40 trials to have encountered one motorcycle travelling towards them and one travelling away from them. During the same 40 trials they would also expect to encounter two pedal cycles and 36 cars. In the memory tests we oversampled bikes, so that of the 24 test locations (12 towards and 12 away), for 6 the correct answer was car, for 3 it was motorcycle, for 3 it was pedal cycle, and for the remaining 12 it was that there was no vehicle present.

3.1.4. Procedure

The procedure was identical to experiment 1 except that before the second block half the participants were asked to say the word "Bike" aloud every time a pedal cycle or motorcycle appeared on the screen. If participants failed to appropriately say "Bike" at any subsequent point during the experiment they were reminded by the experimenter to keep performing this task for the remainder of the experiment.

For experiment 2 memory tests were also categorised and counterbalanced in terms of load, where this was defined by the total number of vehicles present across the three relevant screens. In low load situations there were no more than two vehicles across the three relevant screens, while high load tests were cases where between three and five vehicles were present over the three screens. Load (2 levels) and delay (3 levels) were fully balanced for each participant, while other factors such as screen location (Left, Ahead, Right), the types of vehicles on specific screens, and the order in which specific memory tests were conducted were counterbalanced across participants.

3.2. Results

3.2.1. Wait/Cross decisions

On average drivers missed 6.6 crossing opportunities (range 0 to 33) and made dangerous crossing decisions on 33.4 occasions (range 0 to 100). Overall this performance equates to an average of 90.7% correct decisions (range 76.2 to 99.8). An initial mixed Analysis of Variance with factors of Block (control vs. experimental) and Secondary Task (say bike vs. control), revealed no significant main effects and no interaction between the two factors, [$F(1,30) = 0.659$. $\varepsilon^2_{(p)} = 0.022$, $p = .423$]. Adding the factor of Error Type revealed that there was no overall difference between the frequency of missed opportunities (0.122) and unsafe crossing (0.083) when each was expressed as a proportion of the available opportunities [$F(1,30) = 1.252$, $\varepsilon^2_{(p)} = 0.040$, $p = .272$], but there was an interaction between Block and Error Type, $F(1,30) = 16.721$, $\varepsilon^2_{(p)} = 0.358$, $p < .001$ with missed opportunities specifically becoming more common over time (first block mean = 0.088, second block mean = 0.155), however this was not affected by secondary task [$F(1,30) = 0.678$, $\varepsilon^2_{(p)} = 0.022$, $p = .417$].

For some occasions that a dangerous crossing decision was made there was only one oncoming vehicle on the current three screens. For these cases we can calculate an individual dangerous crossing rate but for pedal cycle and motorcycle screens these screens are relatively rare, so they have been combined into a single category of bikes. This proportion was analysed in a 2x2x3x2 mixed ANOVA with within participants factors of Vehicle Type (car vs. bike), Block (control vs. experimental) and Screen (current, n-1 and n-2) and the between subjects factor of Secondary Task. There was a main effect of Block with dangerous crossings reducing over the two blocks, $F(1,30) = 5.769$, $\varepsilon^2_{(p)} = 0.161$, $p = .023$. The main effect of Screen was significant $F(1.204,36.112) = 20.555$, $\varepsilon^2_{(p)} = 0.407$, $p < .001$. Helmert contrasts revealed that dangerous crossings with a vehicle on the current screen were rarer than those to screens held in memory, $F(1,30) = 22.808$, $\varepsilon^2_{(p)} = 0.432$, $p < .001$, and that dangerous crossings were more common to vehicles viewed two screens previously than one screen ago, $F(1,30) = 16.858$, $\varepsilon^2_{(p)} = 0.360$, $p < .001$. Although there was a marginal effect of Vehicle Type [$F(1,30) = 2.949$, $\varepsilon^2_{(p)} = 0.090$, $p = .096$] with false alarms for bikes being slightly more frequent than those for cars, no other main effects or interactions approached significance.

For comparison with Experiment 1, we conducted a mixed 2x3x2 ANOVA with within participants factors of Vehicle Type (car vs. bike) and Screen (current, n-1 and n-2), and the between participants factor of Experiment (experiment 1 vs experiment 2) which revealed that the interaction between Experiment and Vehicle Type was significant, $F(1,66) = 4.954$, $\varepsilon^2_{(p)} = 0.070$, $p = .029$, with bikes being associated with more dangerous crossings than cars in experiment 2, but fewer dangerous crossings than cars in experiment 1.

In Experiment 2 we have additionally created a subset of balanced safe cross opportunities and dangerous crossing opportunities at high and low levels of load. In all these specific cases the current screen would be empty, but a high load scenario would have two (irrelevant) vehicles travelling away from the junction on the previous two screens, while a low load scenario would have no vehicles travelling away on any of the screens. A 2x2 ANOVA with Error Type (miss vs. false alarm) and Load (low vs. high) revealed a significant interaction between the factors, $F(1,31) = 8.466$, $\varepsilon^2_{(p)} = 0.215$, $p = .007$, such that increased load was associated with more missed opportunities (0.08 vs. 0.19) but not more dangerous crossings (0.23 vs. 0.20).

3.2.2. Memory dependent Variables:

An initial analysis was conducted on the proportion of incorrect responses to the question "was there a vehicle present?" in a repeated measures 2x2x3 design with factors of Direction (towards vs. away), Memory load (low vs. high) and delay (n-1, n-2, n-3). This revealed a main effect of Direction, with memory being better on towards trials, $F(1,31) = 5.936$, $\varepsilon^2_{(p)} = 0.161$, $p = .021$ and Memory Load, with memory being better on low load trials, $F(1,30) = 15.140$, $\varepsilon^2_{(p)} = 0.328$, $p < .001$, but no main effect of Delay [$F(1.786,55.377) = 1.269$, $\varepsilon^2_{(p)} = 0.039$, $p = .287$] and no interactions between the factors. The importance of load rather than delay is plotted in Fig. 6(a). To check whether this was coming separately from vehicle present and vehicle absent trials we have added Error Type (miss vs. false alarm) as an additional factor – doing so reveals a main effect of Error Type, $F(1,31) = 5.215$, $\varepsilon^2_{(p)} = 0.144$, $p = .029$ and an interaction between Error Type and Load, $F(1,31) = 9.448$, $\varepsilon^2_{(p)} = 0.234$, $p = .004$, see Fig. 6(b), but there were no other significant main effects of interactions.

To see whether the say bike instructions have any effect on memory for specific vehicles we have limited the analysis to occasions when a vehicle was actually present and conducted a 3x2x2 mixed design with Vehicle Type (car, motorcycle, or pedal cycle) and Direction (travelling towards the junction, or away from the junction) as within participants factors and Secondary Task (say bike vs. none) as a between participants factor. Here our dependent variable is the proportion of errors – in this case misses. There was a main effect of Direction, $F(1,30) = 8.266$, $\varepsilon^2_{(p)} = 0.216$, $p = .007$, with people missing vehicles on 24.7% of away trials, but only 12.5% of towards trials, but no other main effects or interactions were significant.

As in experiment 1 we can compare memory for specific vehicle types after removing participants who never made a particular response type. This reveals that the accuracy for motorbikes and pedal cycles does not differ, $t(28) = 0.760$, $d = 0.141$, $p = .453$. With pedal cycles and motorbikes combined and five participants with absent responses removed we can conduct a 2x2x2 mixed ANOVA with within participants factors of Vehicle Type (car vs. bike) and Direction (towards vs. away) and the between subjects factor of Secondary Task. This analysis confirms that memory is better for bikes than cars, $F(1,25) = 14.438$, $\varepsilon^2_{(p)} = 0.366$, $p = .001$, and for vehicles travelling towards the junction rather than away, $F(1,25) = 9.626$, $\varepsilon^2_{(p)} = 0.278$, $p = .005$, but no other main effects or interaction were significant.

To compare performance with Experiment 1 without removing any participants' data, we averaged over Direction and types of bike

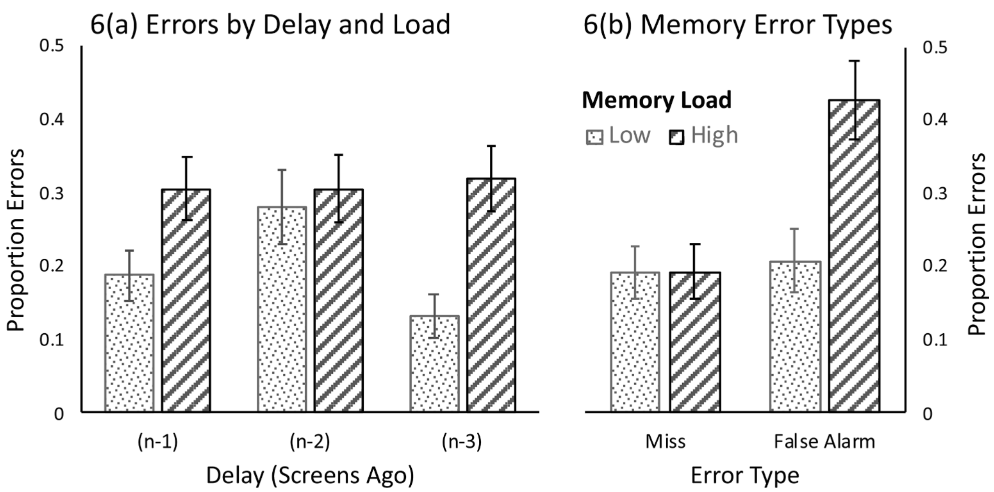

Fig. 6. Memory performance in Experiment 2 as a function of load (low load if there were fewer than 3 vehicles present in the 5 other locations, high load if there were more than 3). Data in panel (a) are split by whether the screen being tested was the previous one (n-1), the one before that (n-2), or three screens ago (n-3). The right side of the figure (panel b) gives data split by whether the error was a miss or a false alarm. Error bars show one standard error above and below each mean.

to conduct a 2x2 mixed ANOVA with factors of Experiment (1 vs. 2) and Vehicle Type (car vs. bike). In addition to the main effect of Vehicle Type, F(1,66) = 22.682, $\varepsilon^2_{(p)} = 0.256$, $p < .001$, there was a marginal interaction between Vehicle Type and Experiment [F (1,66) = 3.944, $\varepsilon^2_{(p)} = 0.056$, $p = .051$], see Fig. 7(a).

For Experiment 2 the overall probability that people were able to correctly remember any specific vehicle that was present and identify its type was 0.68. A comparison with Experiment 1 shows that in addition to the main effect of direction, F(1,66) = 66.906, $\varepsilon^2_{(p)} = 0.503$, $p < .001$, overall performance using this measure is significantly better in experiment 2, F(1,66) = 6.593, $\varepsilon^2_{(p)} = 0.091$, $p = .013$, than it was in experiment 1 (0.56) – Fig. 7(b) shows this for the two experiments with the Experiment 1 data for people additionally remembering the colour of a vehicle added.

3.3. Discussion of experiment 2

Experiment 2 successfully replicated the basic results from experiment 1, both in terms of the finding that unsafe crossing errors come from memory errors, particularly at (n-2), and in the finding that visual short term memory in this task is very poor, and not particularly dependent on delay. Experiment 2 additionally allowed us to explore the role of interference from irrelevant vehicles in this task. In terms of crossing decisions, a high load in terms of vehicles travelling away makes performance worse primarily by increasing the number of missed cross opportunities. This is consistent with the idea that people are generally more cautious when there are more vehicles around, however, while this caution does reduce the number of crossing decisions overall it does not seem to reduce the number of unsafe crossing decisions. One possibility is that in unsafe situations interference from irrelevant vehicles can create both caution and confusion – drivers may know there are several vehicles present but be increasingly unsure as to whether the vehicles are relevant. Load certainly had a dramatic effect on memory for individual vehicles. This effect of load was much larger ($\varepsilon^2_{(p)} = 0.328$) that the other significant effect, that of direction (with memory being better for vehicles travelling toward the junction than away from it, $\varepsilon^2_{(p)} = 0.161$) and the non-significant influence of delay ($\varepsilon^2_{(p)} = 0.039$). While the effect of load was expected – the task feels much more difficult with multiple vehicles present – the finding that it overwhelms any effect of delay was not expected – see Fig. 6(a). Theoretically it is consistent with the idea that load provides two sorts of interference in memory. Proactive interference comes from the fact that when two previous trials, (n-3) and (n-2), with multiple vehicles on them have to be remembered it is hard to remember the last screen looked at (n-1). Retroactive interference comes from the finding that when the last two screens, (n-1) and (n-2) have multiple vehicles on them memory for the oldest screen (n-3) is also impaired. Here it is important to note that load in the cross/wait task referred just to the number of irrelevant vehicles (those travelling away from the junction). In contrast the load measure in the memory test is a combined count of both vehicles travelling towards and away from the junction. Practically this suggests that load caused by items that are deliberately attended to may be more dangerous than background load from unattended items. In the current paradigm is not possible to separate out the effects of attended and unattended items on dangerous crossing decisions because remembering a relevant item should always prevent drivers from crossing. In the real world, however, a driver may attend to many items that are not relevant to their main task and it is easy to see how such divided attention would impair a decision making process that relies on a limited capacity memory system.

Results from cross/wait task additionally highlighted an increase in crossing errors in front of bikes rather than cars that had not been observed in experiment 1. This may be related to the fact that the frequency of bikes in experiment 2 was much lower than that in experiment 1, however, it should be noted that while the interaction between the two experiments and vehicle type was significant, the size of this effect in terms of increased dangerous crossings around bikes was still relatively small ($\varepsilon^2_{(p)} = 0.090$) particularly when compared to the dramatically increased risk to motorcyclists that has been observed in real world crashes. Nonetheless, these results to

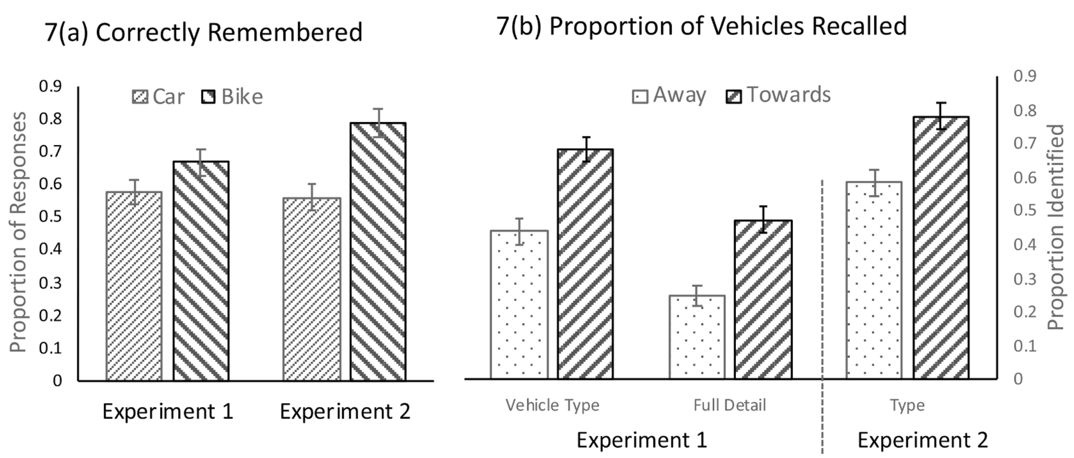

Fig. 7. A comparison of memory for vehicles in Experiments 1 and 2. The left panel (a) shows the proportion of responses correct [hits/(hits + false alarms)] for cars and bikes. The right panel shows the probability that a present vehicle will be remembered and its type accurately identified, and for Experiment 1 the probability that its type and colour (full detail) will be correctly remembered. Error bars show one standard error above and below each mean.

support the general finding that low frequency items in this context are less likely to be perceived than higher frequency ones (e.g. Pammer et al., 2018). The fact that the increase in dangerous crossings in front of bikes in experiment 2 does not interact with delay is consistent with the idea that it arises from an initial perceptual difference at the time the screen is being initially attended to. In contrast it should also be noted that the lower frequency of bikes actually seemed to be associated with improved memory for them in experiment 2 relative to cars - see Fig. 7(a). This is consistent with the general memory findings that rare items (once attended to) are generally well remembered – sometimes referred to as the von Restorff effect (von Restorff, 1933; Reed, 1995).

The say bike, say bike instructions proved to have almost no effect on performance in the cross/wait task. This was unexpected, but it is consistent with the finding that it had almost no effect on memory for bikes. Our hypothesis had been that having the word “bike” in the phonological loop should have supplemented visuospatial memory and improved memory specifically for bikes. We had certainly expected that if a participant had just said “bike” they would be less likely to forget it. However, our results did not seem to support this simple hypothesis. On debriefing participants they often mentioned that they always seemed to be saying bike – since they responded both to bikes travelling both towards them and going away from them, they did not feel that the word “bike” was particularly relevant to the task. With hindsight this does feel like an important detail – in future research we plan to limit verbalisations to oncoming vehicles. It may be that halving the number of times they said “bike” could increase both its specificity and relevance. There is also an issue with our memory tests’ ability to capture information held in the phonological loop. Since we carefully counterbalanced the order of memory tests and randomised whether tests for vehicles towards or away were given first, people would often have read several screens of relevant instructions before actually being asked the cued recall instructions about a specific oncoming bike. Since phonological encoding of even visual information is thought to be automatic (e.g. Baddeley, 1986) it is possible that a mixture of output interference and interference from reading the instructions had overwritten any material in the phonological loop before it could be used in the memory task. Although this counterbalancing was important for our investigation of visual memory it may be that if we want to test phonological memory in this kind of task an alternative memory testing strategy is required.

4. General discussion

The PRC model was designed to account for surprising memory failures in a simulator experiment by Robbins et al., (2019). Although the Robbins et al. experiment had the advantage of being conducted in an advanced driving simulator and requiring participants to actually drive through multiple junctions it meant that research was expensive and time-consuming, and this severely limited the number of trials and memory tests that could be conducted. The primary aim of the experiments in the current paper was to develop a part-task paradigm that could replicate many aspects of the junction crossing task while allowing much more direct tests of visuospatial memory and its role in safe decisions at junctions. The most surprising aspect of the wait/cross task used in both experiments is just how difficult it is to do. Intuitively it feels as if remembering that an oncoming vehicle is present for just two head movements should not be difficult, however, that intuition changes very quickly when participants start doing the cross/wait task. In fact participants missed safe opportunities to cross on about 10% of occasions, and erroneously chose to cross when there was an oncoming vehicle present on about 8% occasions. The task proves remarkably difficult to do in practice and feels as if it is applying a very high memory load. This may be suprising given that people only have to hold one relevant oncoming vehicle in memory to do the task, but in reality the challenge is the fact that memory has to be constantly updated. Previous vehicles have to be forgotten when they are no longer relevant and information has to be updated with every head movement. This is of course characteristic of some real world driving environments – if no other vehicles moved the driving task would be extremely easy. It does raise questions about the validity of our paradigm as a part-task simulation – if drivers on real roads accidentally pulled out into the path of an oncoming vehicle on 8% of occasions the world’s roads would soon consist of nothing but crashed vehicles. Arguably what makes our task so hard is the need to maintain performance at this high memory load over hundreds of trials. Many of the errors made by our drivers could be attributed to the difficulty of maintaining performance despite an array of external and internal distractions (e.g. task unrelated thoughts – Chapman, Ismail, & Underwood, 1999). Drivers at actual junctions will generally be able to limit the time they are performing an equivalent task so internal distractions will hopefully be rare. The problem with external distractions is that drivers may not realise how damaging attention to other things can be to representations that are held in temporary storage. Although the PRC model provides potential mechanisms for improving memory, it also provides a wealth of potential ways in which memory can be interfered with. We have already seen in experiment 2 the way in which attending to other relevant items impairs your memory, but it is easy to see how other visual or auditory distractions could have similar effects.

The idea that retaining relevant information in memory is critical to safe crossing is made abundantly clear by the data presented in Fig. 4. Drivers very rarely made errors when they were looking at an oncoming vehicle, the vast majority of errors come from oncoming vehicles that needed to be held in memory over one and particularly two head movements. People often feel that they must have good memory for the environment around them and assume that if they have looked at something they will remember it – at least for a few seconds. Demonstrations of change blindness make it clear that this is not true and have been proposed as road safety interventions (e.g. Gunnell et al., 2019). The memory tests conducted in both experiments also demonstrate clearly the limitations of visual memory for our driving environment. The data presented in Fig. 7(a) shows dramatically that when people reported remembering a car they were right less than 60% of the time. Although memory is better for relevant vehicles than irrelevant ones, from Fig. 7(b) we can see that in both experiments people were only able to remember the identity of an oncoming vehicle approximately 75% of the time. In experiment 1 where we queried them about the colour of vehicles we found that people were able to remember the type and colour of a vehicle they had just viewed less than 50% of the time if it was travelling towards them, and just 25% of the time if it was travelling away from them. These results represent a dramatic demonstration that the capacity of visuospatial memory is not sufficient to maintain an accurate representation of the view at a junction even when just three vehicles are present on average. It should be noted that we have observed these limitations even with undergraduate students as our participants. We would expect limitations in short term memory to be more dramatic with older drivers, and it is this group who appear to be particularly over-involved in junction crashes (Clarke et al., 2007), it would thus be of great interest to explore these tasks with older drivers in the future. It should be noted, however, that the need to make continuous large head movements may make this task more arduous for older drivers than it was for our student drivers (Herriotts, 2005).

One major limitation of the current task is that it is one that has been developed as a part-task simulation to specifically investigate the role of short-term memory at junctions. In order to obtain precise experimental control, no attempt has been made to simulate a range of other factors that are likely to be relevant at real junctions. The first and most dramatic limitation is the lack of motion cues in this task – at real junctions the precise location and relative motion of the driver's own vehicle and oncoming traffic is of primary importance in determining whether a gap is safe to take. More importantly, the driver has to predict the future motion and locations of all vehicles and potentially update this information in real time even while looking in the opposite direction. In many situations predicting motion requires drivers to predict the intentions of other road users based on extremely subtle cues (e.g. Lee & Sheppard, 2016). The decisions required in the current tasks were simple and unambiguous, which makes it all the more striking that so many errors were observed. Of course we have made the workload high – only one trial in eight is safe to cross, but this is not necessarily unrealistic – at busy times there are many junctions where a driver might to have to make more than eight head movements in each direction before being able to cross. Some of the other aspects of the task we have ignored are the auditory cues that might be provided by an oncoming vehicle, the precise contrast of vehicles, and their variability in appearance, and of course, specific interventions such as daytime running lights that may improve the visibility, and perhaps memorability of motorcyclists. Some of these factors may be important in explaining why memory errors don't create even more real crashes.

One way to get around these problems with the limitations of human memory is to rely on technology, and vehicles are increasingly being fitted with sensors that are not limited by the short term memory constraints imposed by human memory. It may be that at some point most vehicles will make decisions at junctions for their drivers and human limitations will become irrelevant. However, it is likely that for many years a substantial proportion of vehicles throughout the world will be driven by humans and it is important to explore road safety solutions that are available to everyone, not just the richest nations and drivers. A simple awareness of the limitations of visual memory may encourage more caution in situations where information needs to be retained across multiple head movements (Gunnell et al., 2019). We remain optimistic that the phonological strategies implied by the PRC model may be useful, and despite the lack of support from experiment 2 we are continuing to develop and test a more targeted "see bike, say bike" intervention as a way of improving drivers' behaviour around motorcycles at junctions.

Although we found some support for the idea that crashes with motorcycles might be related to motorcyclists' relative infrequency on some roads (Pammer et al., 2018; Wolfe et al., 2005) our findings suggest that this can only be part of the explanation for their high involvement in junction crashes. The research in the current paper provides strong support for the idea of "saw but forgot" crashes (Robbins et al., 2019) both involving cars, motorcycles and pedal cycles. While we did find occasions where drivers would cross in front of oncoming vehicles that they were looking at, these were extremely rare compared to the much more frequent errors associated with memory failures. We suggest that practitioners in many areas of road safety need to be aware that drivers' short term memory can easily be overloaded. This is particularly important given that drivers may not themselves realize that minor distractions can cause them to forget the task-related information they have previously processed. Future work needs to continue to explore whether motorcyclists are indeed more likely to be forgotten by drivers once fixated and to develop interventions that will improve their safety at junctions.

5. Author note

We thank the following University of Nottingham students for their help with some of the stimulus development and pilot testing: John Kiernan, Nikoletta de Scossa, Amy Walsh, Charlie-Luke O'Connor, Amy Monjack, Jayden Gough, Madeline Norman. We would also like to thank Karl Miller for his assistance with the NITES driving simulator, for theoretical development, and for feedback on early versions of this article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Peter Chapman: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – review & editing, Visualization, Supervision. **Seda Orhan:** Investigation, Resources, Data curation, Writing – original draft. **Lily Moore:** Investigation, Resources, Data curation, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Baddeley, A. (1986). *Working memory*. Oxford: OUP.
Baddeley, A. (2010). Working memory. *Current Biology, 20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014
Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422
Baddeley, A., Banse, R., Huang, Y. M., & Page, M. (2012). Working memory and emotion: Detecting the hedonic detector. *Journal of Cognitive Psychology, 24*(1), 6–16. https://doi.org/10.1080/20445911.2011.613820
Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1
Brown, I. D. (2005). *Review of the Looked But Failed to See Accident Causation Factor* - Road Safety Research Report S. No. 60. Department for Transport, U.K.
Chapman, P. R., Ismail, R., & Underwood, G. (1999). Waking up at the wheel: Accidents, attention and the time-gap experience. In A. G. Gale (Ed.), *Vision in Vehicles 7* (pp. 131–138). Oxford: Elsevier.
Clarke, D. D., Ward, P., Bartle, C., & Truman, W. (2007). The role of motorcyclist and other driver behaviour in two types of serious accident in the UK. *Accident Analysis & Prevention, 39*(5), 974–981. https://doi.org/10.1016/j.aap.2007.01.002
Crundall, D., Crundall, E., Clarke, D., & Shahar, A. (2012). Why do car drivers fail to give way to motorcycles at t-junctions? *Accident Analysis & Prevention, 44*(1), 88–96. https://doi.org/10.1016/j.aap.2010.10.017
Crundall, D., Howard, A., & Young, A. (2017). Perceptual training to increase drivers' ability to spot motorcycles at T-junctions. *Transportation Research Part F, 48*, 1–12. https://doi.org/10.1016/j.trf.2017.05.003
Davoodi, S. R., & Hossayni, S. M. (2015). Role of Motorcycle Running Lights in Reducing Motorcycle Crashes during Daytime; A Review of the Current Literature. *Bulletin of Emergency and Trauma, 3*(3), 73–78. https://doi.org/10.7508/beat.2015.03.001
Department for Transport. (2021). *Reported road causalities in Great Britain: motorcycle factsheet, 2020.* https://www.gov.uk/government/statistics/reported-road-casualties- great-britain-motorcyclist-factsheet-2020/reported-road-casualties-in-great-britain- motorcycle-factsheet-2020.
Department for Transport. (2022). *Reported road causalities in Great Britain: motorcycle factsheet, 2020.* https://www.gov.uk/government/statistics/reported-road-casualties- great-britain-motorcyclist-factsheet-2020/reported-road-casualties-in-great-britain- motorcycle-factsheet-2020.
European Commission (2020) Facts and Figures Motorcyclists and moped riders. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. https://road-safety.transport.ec.europa.eu/system/files/2021-07/facts_figures_p2w_final_20210323.pdf.
Fagot, C., & Pashler, H. (1995). Repetition blindness: Perception or memory failure? *Journal of experimental psychology: Human perception and performance., 21*(2), 275.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191, 10.358/BF03193146.
Gunnell, D. O. A., Kunar, M. A., Norman, D. G., & Watson, D. G. (2019). The hazards of perception: Evaluating a change blindness demonstration within a real-world driver education course. *Cognitive Research: Principles and Implications, 4*, 15. https://doi.org/10.1186/s41235-019-0165-4
Herriotts, P. (2005). Identification of vehicle design requirements for older drivers. *Applied Ergonomics, 36*, 255–256. https://doi.org/10.1016/j.apergo.2005.01.002
Herslund, M. B., & Jørgensen, N. O. (2003). Looked-but-failed-to-see-errors in traffic. *Accident Analysis & Prevention, 35*(6), 885–891. https://doi.org/10.1016/s0001-
Horswill, M. S., Helman, S., Ardiles, P., & Wann, J. P. (2005). Motorcycle accident risk could be inflated by a time to arrival illusion. *Optometry and Vision Science, 82*(8), 740–746. https://doi.org/10.1097/01.opx.0000175563.21423.50
Horowitz T. S., & Wolfe J. M. (1998). Visual search has no memory. Nature. 1998 Aug 6;394(6693):575-7. doi: 10.1038/29068. PMID: 9707117.
Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior, 1*(3), 153–161. https://doi.org/10.1016/S0022-5371(62)80023-1
Keyes, H., Green, F., Compton, C., & Staton, M. (2019). Short-term cognitive conspicuity training does not improve driver detection of motorcycles at road junctions: A reply to Crundall, Howard, and Young (2017). *Transportation Research Part F: Traffic*.
Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology., 55*(4), 352–358. https://doi.org/10.1037/h0043688
Lee, Y. M., & Sheppard, E. (2016). The effect of motion and signalling on drivers' ability to predict intentions of other road users. *Accident Analysis & Prevention, 95*(Pt A), 202–208. https://doi.org/10.1016/j.aap.2016.07.011
Lee, Y. M., Sheppard, E., & Crundall, D. (2015). Cross-cultural effects on the perception and appraisal of approaching motorcycles at junctions. *Transportation Research Part F: Traffic Psychology and Behaviour, 31*, 77–86. https://doi.org/10.1016/j.trf.2015.03.013
Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience, 13, Article 57*. https://doi.org/10.3389/fnhum.2019.00057
Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science*. 2003 Oct;12(5):180-4.
Miller, K. A., Chapman, P., & Sheppard, E. (2021). A Cross-Cultural Comparison of Where Drivers Choose to Look When Viewing Driving Scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 639–649. https://doi.org/10.1016/j.trf.2021.07.013
National Safety Council (2022). *NSC Injury Facts.* https://injuryfacts.nsc.org/motor-vehicle/road-users/motorcycles/.
National Travel Survey (2022). *National Travel Survey: 2021.* Updated 31 August 2022. Department for Transport UK: https://www.gov.uk/government/collections/national-travel-survey-statistics.
Olson, P. L., Halstead-Nussloch, R., & Sivak, M. (1981). The effect of improvements in motorcycle/motorcyclist conspicuity on driver behavior. *Human Factors, 23*(2), 237–248. https://doi.org/10.1177/001872088102300211
Pammer, K., Sabadas, S., & Lentern, S. (2018). Allocating attention to detect motorcycles: The role of inattentional blindness. *Human Factors, 60*(1), 5–19. https://doi.org/10.1177/2F0018720817733901
Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017
Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10. https://doi.org/10.3389/neuro.11.010.2008
Reed, H. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review., 2*(1), 105–112. https://doi.org/10.3758/BF03214414
Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368–373.
Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *The Quarterly Journal of Experimental Psychology, 69*(10), 1969–1995. https://doi.org/10.1080/17470218.2014.914546
Robbins, C. J., Allen, H. A., & Chapman, P. (2018). Comparing drivers' gap acceptance for cars and motorcycles at junctions using an adaptive staircase methodology. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 944-954. https://doi.org/10.1016/j.trf.2018.07.023.

Robbins, C. J., Allen, H. A., & Chapman, P. (2019). Comparing Drivers' Visual Attention at Junctions in Real and Simulated Environments. *Applied Ergonomics, 80*, 89–101. https://doi.org/10.1016/j.apergo.2019.05.005

Robbins, C. J., Allen, H. A., Miller, K. A., & Chapman, P. (2019). The 'Saw but Forgot' error: A role for short-term memory failures in understanding junction crashes? *PloS ONE, 14*(9), e0222905.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*(9), 1059–1074. https://doi.org/10.1068/p281059. PMID: 10694957.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261–267. https://doi.org/10.1016/S1364-6613(97)01080-2

Simpson, M. J. (2017). Mini-review: Far peripheral vision. *Vision Research, 140*, 96–105. https://doi.org/10.1016/j.visres.2017.08.001

Solís-Marcos, I., & Kircher, K. (2019). Event-related potentials as indices of mental workload while using an in-vehicle information system. *Cogn Tech Work, 21*, 55–67. https://doi.org/10.1007/s10111-018-0485-z

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied., 74*(11), 1.

Surprenant, A. M., & Neath, I. (2009). *Principles of memory*. New York: Psychology Press.

To, M. P. S., Regan, B. C., Wood, D., & Mollon, J. D. (2011). Vision out of the corner of the eye. *Vision Research, 51*(1), 203–214. https://doi.org/10.1016/j.visres.2010.11.008

von Restorff, H. (1933). "Über die Wirkung von Bereichsbildungen im Spurenfeld" [The effects of field formation in the trace field]. *Psychologische Forschung [Psychological Research] (in German), 18*(1), 299–342. https://doi.org/10.1007/BF02409636

Walton, D., Buchanan, J., & Murray, S. J. (2013). Exploring factors distinguishing car- versus-car from car-versus-motorcycle in intersection crashes. *Transportation Research Part F: Traffic Psychology and Behaviour, 17*, 145–153. https://doi.org/10.1016/j.trf.2012.11.001

White, & Caird, J. K. (2010). The blind date: The effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident Analysis and Prevention, 42*(6), 1822–1830. https://doi.org/10.1016/j.aap.2010.05.003

Wolfe, B., Dobres, J., Rosenholtz, R., & Reimer, B. (2017). More than the Useful Field: Considering peripheral vision in driving. *Applied Ergonomics, 65*, 316–325. https://doi.org/10.1016/j.apergo.2017.07.009

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual search. *Nature, 435*, 439–440.

World Health Organization (2022). Powered two-and three-wheeler safety: a road safety manual for decision- makers and practitioners, second edition. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.