



How does drivers' visual search change as a function of experience? A systematic review and meta-analysis

Chloe Robbins*, Peter Chapman

School of Psychology, University of Nottingham, Nottingham NG7 2RD, United Kingdom



ARTICLE INFO

Keywords:

Driving
Experience
Novice drivers
Experienced drivers
Visual search
Eye movements

ABSTRACT

Novice drivers are statistically over-represented in reported road crashes, with recent evidence suggesting that some of this increased crash involvement may be a result of limitations in their cognitive processing. Such processing has typically been measured by recording drivers' patterns of eye movements, however, the exact ways in which eye movements are reported and interpreted varies substantially between different studies in the literature. Therefore, the objective of this systematic review was to investigate whether novice drivers and experienced drivers do differ in clear and reproducible ways in their visual search.

Studies were identified through searches of Web of Science, Medline, TRID Database, and the TRB Research in Progress Database, with no restrictions on publication status. Studies were included if they compared the visual search of a novice driver group (< 3 years driving experience) and an experienced driver group (> 3 years driving experience) using an eye tracking method and reported at least one of the following four visual search outcomes: fixation durations, horizontal spread of search, vertical spread of search and number of fixations. Two reviewers independently screened searches and assessed the full texts of potentially included studies.

Of the 235 studies initially identified 18 were included in the review, with 13 studies reporting sufficient data to be included in the meta-analysis for at least one outcome measure. Given that the included studies deployed a range of method types, additional sub-group analyses were conducted using this factor. Sensitivity analyses were also conducted by temporarily removing extreme experience groups (e.g. driving instructors and learner drivers) in order to test the effect of different levels of experience and training.

The meta-analyses, along with support from results discussed narratively, revealed that novice drivers have a narrower horizontal spread of search compared to experienced drivers, however, there were no overall differences in fixation durations, vertical spread of search or number of fixations when the studies were pooled together. These findings have important primary implications for the development of novice training interventions, with novice drivers needing to develop a broader horizontal spread of visual search, but not to necessarily learn to fixate further down the road. Subgroup analyses also provided considerations for future research studies in terms of the experience of the driver groups, and the method type used.

1. Introduction

Driving on public roads is a highly complex and responsible task, with mistakes or risk-taking having potentially fatal consequences (Drews et al., 2008). It is widely agreed that it takes time and experience to become a fully safe and competent driver (Mayhew et al., 2003). Given the importance of visual information when driving, it is unsurprising that there have been studies investigating drivers' visual search with a particular focus on experience, dating back more than 40 years (e.g. Mourant and Rockwell, 1972; Renge, 1980). However, although there are many studies investigating this topic, it is often difficult to compare these due to the variety of methodologies deployed

(Cruddall and Underwood, 1998).

1.1. Age and experience

Driving statistics for many years have found that driver age and experience both make independent contributions to high crash rates (McCart et al., 2009). Young car drivers in the UK between the ages of 17 and 25 are statistically over-represented in reported road accidents compared to older drivers aged 25 and above (DfT, 2015). In the UK, young car drivers have been found to make up 18% of all reported road crashes, which is considerably higher than the 5% of miles they account for (DfT, 2015). In the US, young drivers which include drivers between

* Corresponding author.

E-mail address: chloe.robbs@nottingham.ac.uk (C. Robbins).

<https://doi.org/10.1016/j.aap.2019.105266>

Received 26 February 2019; Received in revised form 9 July 2019; Accepted 10 August 2019

Available online 29 August 2019

0001-4575/© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the ages of 15–20 years made up 9% of all fatal crashes in 2016, despite accounting for only 5.4% of all licensed drivers (National Highway Traffic Safety Administration, 2016). Globally, it has been found that road injuries sustained from driving are the leading cause of death for people between the ages of 15–29 years (World Health Organization, 2018).

Driver inexperience is also one of the most frequently reported contributory factors towards traffic crashes in the UK literature (Chapman and Underwood, 1998a) and therefore novice drivers are particularly at risk (Clarke et al., 2006). However, methodologically, it has always been difficult to separate the effects of age from those of experience on accident frequencies, as they are typically closely inter-related (McCartt et al., 2009). Nonetheless, in reviews conducted in the US and Canada, it has been found that increased driving experience has a protective effect on crash risk, with increases in driving experience being associated with reductions in crash rates for drivers of all ages (Mayhew et al., 2003; McCartt et al., 2003). This evidence has helped shape countermeasures, for example, the graduated driver licensing (GDL) scheme, with these programs applying to all new drivers regardless of age (McCartt et al., 2009).

While some studies have used the distance driven since passing a driving test as a measure of experience, many drivers find this information difficult to report accurately. Moreover, raw measures of experience based on distance driven can lead to confusing positive relationships between experience and crash involvement because of basic exposure effects, with individuals who drive more frequently being more likely to be involved in a crash (Peck, 1993). Therefore, the most common independent variable relating to driving experience in published research tends to be length of licensure. Because time since licensure is easy to calculate as a measure of driving experience and can be related directly to crash statistics, the observation has been made in the UK that drivers with less than 3 years driving experience are statistically over-represented in reported crashes compared to drivers with more than 3 years driving experience (Crundall et al., 2003; Clarke et al., 2006).

1.2. 'A failure to look properly'

In the UK, contributory factors associated with traffic crashes are routinely assessed by the completion of a STATS 19 form by police officers at the scene (Haigney, 1995). The police officer is required to provide factors that are thought to have contributed to the crash, with the most common category titled 'Driver Error'. This category includes various types of perceptual errors such as a 'failure to look properly' and 'failure to judge the other person's path or speed' (Sabey and Staughton, 1975). Since 2005, 'failed to look properly' has continued to be the most frequently reported contributory factor for reported road crashes, with 39% of crashes described using this contributory factor in 2017 (DfT, 2018). For this reason, drivers' visual search on the road has been under intensive investigation, with researchers particularly interested in what affects drivers' eye movements on the road (Crundall and Underwood, 1998).

1.3. Previous literature

When a new driver becomes qualified, it is relatively easy to confirm that they possess adequate motor skills to control the vehicle (steering, braking) however, there is evidence to suggest that their higher order cognitions are not fully developed (Isler et al., 2011). Cognitive processing demands are reflected by several aspects of eye movement behaviour, therefore measuring this behaviour is a strong indication of cognitive difficulty on the road (Chapman and Underwood, 1998b; Underwood, 2007). Studies investigating drivers' visual search typically use eye tracking technology, allowing for moment-by-moment tracking of the driver's eye movements over the visual scene (Bremmer et al., 2009). It is typical for drivers' general visual search to be measured

over the visual scene in terms of fixation durations (how long each fixation lasts before the next saccade), the number of fixations made in a given time period, and horizontal and vertical spread of search (in terms of the variance in fixation locations across the visual field).

A previous information processing model for the control of eye movements proposed by Findlay and Walker (1999) provides a theory for deriving predictions about the distribution of fixations in a given scene, by identifying two competing pathways known as the "when" and "where" pathways. The decision to move the eyes is based on the competing demands of a "fixate" centre (which attempts to process information currently available at the point of gaze) and a "move" centre (which identifies potential locations within a broad saliency map to redirect gaze towards). The decision about when to make a new saccade is thus both related to the information that is being processed foveally (with fixation durations often regarded as an indicator of processing load - Cohen, 1981) and the information potentially available from other areas of the visual field. Within this model activation in both pathways is dependent on a mixture of both top down and bottom up factors.

In the context of driving, the model allows us to predict that experienced drivers might be able to process items at the point of gaze faster than novices, and that they would thus show shorter fixation durations overall. Such shorter fixation durations could allow additional visual search to take place that might be reflected in them achieving more fixations overall, or a broader spread of search over the visual field. Top down factors based on driver experience may also influence the "where" pathway, suggesting areas of the visual field for new fixations. Thus, an experienced driver may choose to fixate areas of low visual salience because of the knowledge that they are sources of potential future hazard-related information. The potential interaction between processing at the point of fixation and the processing of peripheral information in a driving context has been demonstrated by Crundall et al. (2002), who found that both experienced drivers' and novice drivers' ability to spot peripheral targets was reduced when a hazard was present at the point of fixation.

The effect of experience on drivers' visual search has been extended by investigating extreme levels of experience and training. Advanced driver groups such as driving instructors and police drivers are of specific interest as they are among the most skilful drivers on the road. Their training heavily relies on improving observation on the road, with both the Road Craft Manual for police drivers (Coyne et al., 2007) and the Driver Instructor Handbook (Miller and Stacey, 2013) stressing the importance of improving scanning of the environment and peripheral vision. On the other hand, learner drivers are also of interest, as this minimal driving experience provides further insight into the role of driving experience on drivers' search on the road (Konstantopoulos et al., 2012).

1.4. The effect of method type on drivers' visual search

Studies investigating this topic have deployed a variety of method types including both simple methodology such as static images and video clips of driving scenes (e.g. Huestegge et al., 2010; Yeung and Wong, 2015), and immersive methodology such as driving simulators and on-road studies (e.g. Bos et al., 2015; Jiang et al., 2012). While images and video clips have been used in many studies due to the practical ease of the method and the ability to expose drivers to multiple driving clips in a short time, there are also some important limitations. Firstly, these methods fail to provide the driver with any element of vehicle control. Although novice drivers are believed to have adequate motor control skills (Deery, 1999), the elimination of this element may free up extra resources, which are usually needed for basic motor control, in order for drivers to scan the environment (Crundall et al., 2012).

In addition, it has also been seen that the visual field which drivers have access to during the driving task can cause differences in drivers'

visual search strategies (Di Stasi et al., 2011; Alberti et al., 2012). When comparing novice and experienced drivers' visual search under a narrow and wide field of view in a driving simulator, it was found that only experienced drivers made use of the wider eccentricities when identifying a hazard, demonstrating a larger difference in horizontal spread of search (Alberti et al., 2014). This finding suggests that the immersiveness of the driving environment may have different effects on novice drivers' and experienced drivers' visual search.

1.5. Place for the systematic review and rationale

Although there has been more than 40 years of studies into visual search and driving expertise, minimal reviews of the literature have been conducted. Green (2007) assessed eye movements while driving, by focusing on the effects of the road environment, driver characteristics and in-vehicle devices on drivers' eye movements. This review highlighted that in an era of driver distraction (e.g. driver information systems) and automated driving (e.g. adaptive cruise control), where these systems have the potential to cause visual interference (Chisholm et al., 2008), reviews of drivers' visual search on the road could provide insights into how these systems might affect driving. In a more recent review, Fisher et al. (2016) narratively summarised a proportion of the existing literature regarding eye movements and driving, focussing on how novice drivers' eye movements differed from experienced drivers on the road.

While both Green (2007) and Fisher et al. (2016) both stressed the importance of reviewing the driving literature regarding the changes in eye movements as a function of experience, there is yet to be a systematic review that seeks to gather all evidence that fits a pre-specified eligibility criterion in order to address the specific question of whether drivers' visual search is related to their level of driving experience. Therefore, this systematic review is the first to assess the relationship between driving experience and drivers' general visual search, synthesising non-randomised controlled studies which compare the visual search of novice and experienced drivers. By using a systematic method and meta-analysis, this is thought to reduce bias and produce reliable findings from which conclusions can be drawn (Antman et al., 1992), as results of independent studies can be statistically summarised (Glass, 1976). Systematic reviews are also extremely important in regards to informing policy decisions, particularly given that both researchers and policy makers are concerned with driver safety.

This systematic review has important primary implications for road safety, particularly in regards to visual search strategy interventions for novice drivers. These interventions include road commentary training for improving search allocation (Crundall et al., 2010; Cantwell et al., 2013; Castro et al., 2016), hazard perception training for anticipating and detecting hazards (Chapman et al., 2002; Horswill et al., 2013) and graduated driver licensing (GDL) schemes in Australia, New Zealand and Canada (Hartling et al., 2004), where specific training to maximise visual search across the driving scene is given. The current systematic review allows for a better understanding of how novice and experienced drivers distribute their visual search across the roadway, and therefore can help inform such interventions.

1.6. Objectives

This systematic review and meta-analysis have been undertaken to investigate whether driving experience relates to drivers' general visual search, by comparing novice drivers' and experienced drivers' fixation durations, spread of search (horizontal spread of search and vertical spread of search) and number of fixations over the driving scene.

Given that studies investigating this topic have deployed a variety of method types, additional sub-group analyses will be conducted by categorising the included studies by method type (simple methodology or immersive methodology). This allows for all studies, irrelevant of choice of method, to be included in the overall meta-analysis of each

outcome measure, as well as investigating the effects of this factor on drivers' visual search.

2. Methods

2.1. Criteria for considering studies for this review

The 27-item PRISMA checklist was used when conducting and reporting this systematic review and meta-analysis (Moher et al., 2009). This systematic review identified all studies which investigated whether driving experience related to drivers' visual search by comparing the visual search of novice and experienced drivers. Since this is a between subject comparison of novice and experienced driver groups from the population, the studies feature no formal randomisation as it is not possible to randomly assign drivers to one of the two groups. The included studies used various methods to investigate potential differences in visual search including on-road studies, driving simulator studies, video recordings and static images of driving scenes.

2.1.1. Inclusion criteria

This review considered all studies that met the inclusion criteria. As there is no single objective defined measure of a novice and experienced driver, we used a practically important distinction based on crash statistics. It is clear that drivers with less than 3 years driving experience are statistically over-represented in reported road crashes compared to drivers with more than 3 years' experience (Clarke et al., 2006), with the reduction in reported road crashes over the first 3 years commonly including those where the driver is primarily at fault (rear end stunts, turns across traffic). Although 3 years is a rather arbitrary cut-off, and may not conform with some literature which suggests that crash risk drops most dramatically during the first 6 months (Mayhew et al., 2003) or first year of driving (Bingham and Shope, 2004), it does provide a clearly defined point around which there is an undeniable reduction in the disproportionate crash involvement associated with novice drivers, and allows for more studies to be considered for inclusion in the systematic review. So, for the current review a novice driver was defined as a driver who had no more than 3 years driving experience after passing the practical test and additionally included learner drivers. In contrast, experienced drivers were defined as those who had 3 or more years driving experience after passing the practical test. Studies that investigated groups of drivers with advanced training such as driving instructors, taxi drivers and police pursuit drivers were included. Only studies that examined the visual search of drivers with normal and corrected normal vision (glasses and contact lenses) were included. Participants in the studies could be of any age or gender. All included studies were published in the English language, though we made no country, date or publication restriction to the search.

2.1.2. Exclusion criteria

The systematic review criterion therefore excluded studies which investigated the visual search of non-drivers. We additionally excluded studies which had passengers present in the car or simulator during the experiment (except the researcher) to eliminate any distraction effects. Studies that only investigated differences in novice and experienced drivers' visual behaviour under the influence of alcohol, drugs, fatigue or in-car distractions such as a mobile phone were also excluded.

2.1.3. Outcome measures

The review considered studies that investigated novices' and experienced drivers' visual search using an eye tracking method and measured at least one of the following outcome measures: fixation durations, spread of search and number of fixations. Fixation durations reflect the length of time drivers generally hold their eyes in each location before moving on. In driving data smooth pursuits on moving objects are typically included as fixations. Spread of search is divided into two measures: horizontal spread of search and vertical spread of

search. These measures show the variance of fixations over the horizontal and vertical axis. Finally, the number of fixations is simply how many fixations a driver made over the driving scene and provides a measure of sampling rate. Although we would generally expect the number of fixations to be inversely proportional to the fixation duration, the ways in which these measures are recorded and reported in individual studies means that this is not always the case. Although some studies provide additional details (e.g. fixations on specific aspects of the road environment), these are not universally provided and there is great variety in the way such details are categorised making them hard to use in a systematic review. Therefore, only studies that reported at least one of these four measures over the whole driving scene, and not in areas of specific interest were included. These general visual search measures were chosen as outcome measures as they are less sensitive to heterogeneity in the methods and tasks used as opposed to capturing the sequences of fixations for hazard anticipation.

2.2. Search methods for identification of studies

The search strategy used was developed to find peer reviewed full journal articles and abstracts (subject to enough information), grey literature including conference proceedings and current ongoing unpublished research.

Electronic searches included Web of science (May, 1900- Jan, 2019), Medline OVID (1946- Jan, 2019) and TRID, the TRIS and ITRD (1990- Jan, 2019) Database. The TRID database is an integrated database that combines the records from TRB's Transportation Research Information Services (TRIS) Database and the International Transport Research Documentation (ITRD) Database, and therefore is a key database of the review. In addition, current ongoing research was also searched electronically using the TRB Research in Progress (RiP) Database (1990- Jan, 2019). These databases were chosen as they are key transport databases and it was unlikely that new studies could have been found elsewhere.

A search strategy was developed to include all relevant keywords relating to drivers, experience, and visual search in each resource. In order for a record to be included in the initial search, the study must have included at least one word or phrase from each of the three categories. See Supplementary File 1 for the search strategy and list of the specific keywords used.

2.3. Data collection and analysis

Following the electronic searches, all citations were downloaded into Mendeley and all duplicates were removed. In regards to the selection of studies, this involved a two-step process. Firstly, the initial search results from the electronic databases were screened against the inclusion criteria by two reviewers who read the titles, abstracts and keywords to identify the studies with potential relevance. Secondly, the full text of the selected citations were obtained and assessed. Two independent reviewers decided on the study's inclusion using the predetermined inclusion criteria. The studies that were not included can be seen in Supplementary File 2, with the reasons for exclusion provided.

The results of the study selection are reported using a PRISMA flow diagram (Moher et al., 2009) in Fig. 1. Qualitative synthesis refers to the number of studies that met the inclusion criteria and therefore could be discussed narratively, and quantitative synthesis refers to the amount of studies that provided the necessary values to be included in the meta-analyses.

2.4. Critical appraisal and data extraction

All studies were critically appraised by two independent reviewers at the study level for methodological quality using the standardised critical appraisal tool, McMaster Critical Appraisal Tool for Quantitative Studies (Law et al., 1998). This tool was chosen due to its

relevance for quantitative, non-randomised controlled studies, therefore had the most fitting criteria for studies included in this review. All papers, regardless of the results of their methodological quality, underwent data extraction and synthesis where possible. The critical appraisal process allowed for full engagement with all of the included papers. The results of the critical appraisal are reported in Tables 1 and 2.

2.4.1. Dealing with missing data

Where data were missing, the authors were contacted and the additional data was requested. If an included study did not report a particular outcome, this study was not included in the analysis of that outcome.

2.5. Data synthesis

2.5.1. Quantitative synthesis

The data from the included studies was synthesised using a meta-analysis where possible, using Review Manager 5 (version 5.1 Nordic Cochrane Centre, the Cochrane Collaboration). There were 13 studies with sufficient data details to calculate the standardised mean difference with 95% confidence intervals for at least one outcome measure. See Table 1.

Standardised mean differences were used as the summary statistic due to the included studies measuring horizontal spread of search, vertical spread of search and number of fixations in a variety of ways. Horizontal and vertical spread of search were measured by the standard deviation of x and y locations in degrees or pixels, however, the available field of view and calibration range of the eye-trackers differed dramatically between studies meaning that raw values could not be directly compared. Number of fixations were measured by the mean number, or total number of fixations during the driving clip, however, these values differed considerably due to the varying length of the driving scenes. Fixation durations were always measured in milliseconds. See Table 1 for each individual studies unit of measurement for each outcome. Given that the remaining measures differed in detail between studies it was necessary to standardise these results to a uniform scale before combining them (Higgins and Green, 2011).

Statistical testing was used (Z, Chi-Square) to investigate the significance of the overall effect for each outcome measure, and for overall subgroup differences (Polanin and Pigott, 2015). As previous reviews have acknowledged a need for consistency in reporting meta-analysis results, these statistical tests will have their corresponding 95% confidence intervals and measure of heterogeneity (I^2).

Heterogeneity was assessed statistically using the standard I^2 test, looking at similarities of studies. While it is acknowledged that determining what constitutes a large I^2 value is subjective, the following rule has been previously suggested (Schünemann et al., 2013). If the heterogeneity is between 0% and 40%, then a fixed effects model should be used whereas if the heterogeneity is between 40% and 85% then a random effects model should be used.

2.5.1.1. Coding of outcome measures. In regards to the coding of outcome measures for each study, values were averaged on occasions where the measure had been separately calculated for different environmental demands (e.g. rural, suburban and dual carriageway; high, medium and low hazardous clips; daytime and night time clips). This approach was taken in order to resolve dependence of the effect sizes in the meta-analysis when multiple measures were available for a single construct (Scammacca et al., 2014). This method is in accordance to Cooper's (1998) shifting-unit-of-analysis approach.

On occasions where a study had both an advanced group of drivers (i.e. police pursuit drivers) and an experienced group of drivers with no additional training, these two groups were integrated into a single experienced group for the overall meta-analysis of each outcome measure, by averaging the values for these drivers. The combining of these driver

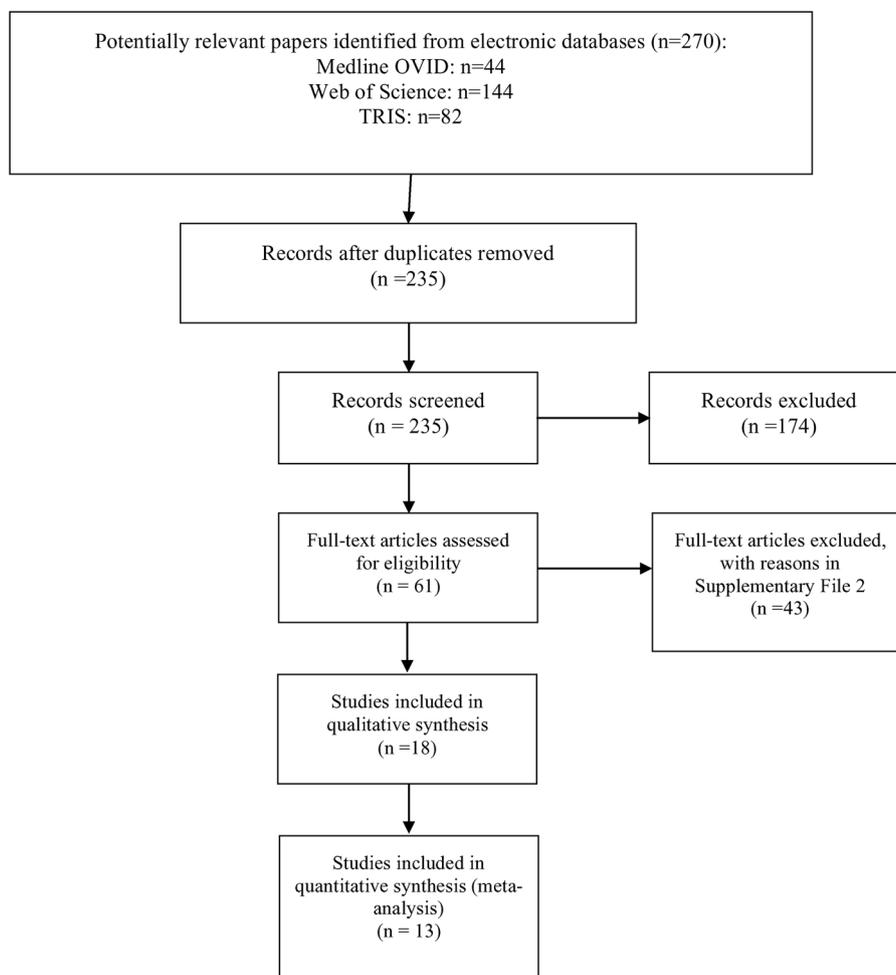


Fig. 1. A PRISMA flow diagram for the number of records included in each stage of the review.

groups meant that all drivers included in the experienced driver group still met the inclusion criteria of more than 3 years driving experience. The inclusion of these advanced driver groups also allowed for the sensitivity analysis detailed below.

2.5.2. Qualitative synthesis

Five studies did not report sufficient data to calculate a mean difference and 95% confidence intervals (Bos et al., 2015; Crundall et al., 1999; Laya, 1992; Lehtonen et al., 2014; Mourant and Rockwell, 1972) and therefore given that statistical pooling was not possible, these findings will be discussed narratively. This qualitative synthesis allows for the findings of these studies to still be integrated in the review as they met the pre-defined inclusion criteria (Ryan, 2013), See Table 2.

2.5.3. Subgroup analysis

Due to the variety of method types used to compare experienced and novice drivers' visual search, subgroup analyses were conducted as there was sufficient data to determine whether outcome measures vary according to method type (Fu et al., 2011), which was either simple methodology or immersive methodology. Simple methodology is defined as a method that presents the visual driving scene on a screen that subtends less than 90 degrees of visual angle, and where the participant has to merely observe the driving scene, requiring no form of vehicle control. These methods usually involve drivers watching static images or video clips. Immersive methodology is defined as a method that presents the visual driving scene on a screen that subtends at least 90 degrees of visual angle, and requires the participant to control a vehicle throughout the driving task. The most common forms of method are

driving simulators (both medium and high fidelity) and on-road studies. Each included study fell clearly into one of the categories as defined above (simple vs. immersive). The coding of these outcomes as a function of method was the same as those previously stated, by averaging across the different road demands in each study.

2.5.4. Sensitivity analysis

To examine the effect of removing studies with the greatest potential risk of bias, a sensitivity analysis was conducted where necessary to test decisions made regarding the inclusion of learner drivers and experienced driver groups who have additional driver training. The main overall analysis for each outcome measure was repeated with these studies temporarily removed.

In addition, given that the definitions of a novice and experienced driver varies considerably across studies, the definition of novice drivers as those with less than 3 years of licensure, and experienced drivers being those with more experience than this is a potentially controversial one. Other studies have defined novice drivers as having held a licence for less than a year (Bingham and Shope, 2004) and experienced drivers as having held a licence for at least 5 years (Chapman and Underwood, 1998a, 1998b). Therefore, the main analysis for each outcome measure was also repeated by removing the studies that included novice drivers with more than 1-year's experience, and experienced drivers with less than 5-years' experience, making the eligibility criteria for novice and experienced drivers more restrictive. The removal of these studies reduced the number of studies that were included in the meta-analysis for each measure but did not change the overall effect for each outcome measure, see S3 for the full restricted

Table 1
Shows the 13 studies that were included in the meta-analysis for at least one measure, as well as the information collected for each of these studies during the critical appraisal.

Study	Method	Task Type	No of Exp	Average Experience (m)	Extra Experienced Group	No of Nov	Average Experience (m)	Learner Drivers	Outcome Reported in Meta-Analysis	Units of Outcome Measures	Outcomes Reported not in Meta-analysis
Alberti et al. (2014)	Simulator	Hazardous Situations	20	126	none	20	19	none	Horizontal Spread of Search	Degrees	None
Borowsky and Oron-Glad (2013)	Video Clips	Hazardous Situations	27	91.1	Includes 10 taxi drivers	17	1.5	none	Vertical Spread of Search	Degrees	No of Fixations and Horizontal Spread of Search
Chapman and Underwood (1998a)	Video Clips	Rural, Suburban, Urban	26	84	none	51	3	none	Mean Fixation Duration, Vertical Spread of Search	Milliseconds, Degrees	Horizontal Spread of Search
Crundall and Underwood (1998)	On-road	Rural, Suburban, Urban	16	108	none	16	2	none	Mean Fixation Duration, No of Fixations, Horizontal/Vertical Spread of Search	Milliseconds, Total Number, Degrees	None
Crundall et al. (2003)	Video Clips	Hazardous Situations	32	226.79	Includes 16 Police pursuit drivers	16	34.8	none	Mean Fixation Duration, Horizontal Spread of Search	Milliseconds, Pixels	Vertical Spread of Search
Hills et al. (2018)	Static Images	Hazardous Situations	20	99	none	36	12	none	Horizontal Spread of Search	Proportion of Eye Movements	None
Huestegge et al. (2010)	Static Images	Hazardous Situations	20	60	none	20	9	Includes learner drivers	Mean Fixation Duration, No of Fixations	Milliseconds, Mean Number	None
Jiang et al. (2012)	On-road	Rural, Urban, Dual Carriageway	5	139.19	none	5	36	none	Mean Fixation Duration, Horizontal/Vertical Spread of Search	Milliseconds, Degrees	None
Kahana-Levy et al. (2019)	Video Clips	Hazardous Situations	35	102.95	none	23	6.45	none	Horizontal/Vertical Spread of Search, No of Fixations	Degrees	None
Konstantopoulos et al. (2010)	Simulator	Differing Visibility	10	408	All driving instructors	11	0	All learner drivers	Mean Fixation Duration, No of Fixations, Horizontal/Vertical Spread of Search	Milliseconds, Total Number, Degrees	None
Underwood et al. (2002a)	Video Clips	Rural, Urban, Dual Carriageway	16	55.07	none	16	1	none	Mean Fixation Duration, Horizontal/Vertical Spread of Search	Milliseconds, Degrees	None
Underwood et al. (2003)	On-road	Rural, Urban, Dual Carriageway	16	108	none	20	2.4	none	No of Fixations	Mean Number	None
Yeung and Wong (2015)	Video Clips	Hazardous Situations	28	99.4	none	14	24	none	Mean Fixation Duration	Milliseconds	Horizontal Spread of Search

Table 2 Shows the 5 studies that were not included in the meta-analysis, as well as the information collected for each of these studies during the critical appraisal.

Study	Method	Task Type	No of Exp	Average Experience (m)	Extra Experienced Group	No of Nov	Average Experience (m)	Learner Drivers	Outcome/s Measured	Reason for exclusion in Meta-analysis
Bos et al. (2015)	Simulator	Hazardous Situations	8	162	none	10	30	none	No of Fixations, Horizontal/Vertical Spread of Search	Means and SD of experienced and novice drivers not reported
Crundall et al. (1999)	Video Clips	Hazardous Situations	20	60	none	20	2.5	none	Mean Fixation Duration, Horizontal/Vertical Spread of Search	Means and SD of experienced and novice drivers not reported
Laya (1992)	Simulator/On-road	Driving around curves	8	60	none	8	12	none	Mean Fixation Duration	Means and SD of experienced and novice drivers not reported
Lehtonen et al. (2014)	On-road	Driving around curves	9	120	none	9	18	none	Horizontal/Vertical Spread of Search	Means and SD of experienced and novice drivers not reported
Mourant and Rockwell (1972)	On-road	Differing Road Types	4	60	none	6	0	All learner drivers	Horizontal/Vertical Spread of Search	Means and SD of experienced and novice drivers not reported

analysis.

It must also be noted that a large proportion of the studies included in the meta-analyses were conducted at the University of Nottingham, with many of the same researchers present. Due to this, the participants that took part in the on-road study by Crundall and Underwood (1998) were a subset of the participants used in the hazard perception clip study by Chapman and Underwood (1998a). The second author of the current paper is also a co-author on a substantial proportion of the included studies, therefore for this reason, the inclusion and appraisal of studies in the review was conducted solely by the first author and a second independent reviewer. An additional sensitivity analysis was conducted to remove any studies that have the same co-author as the current paper. Again, the removal of these studies did not change the overall effect for each outcome measure, see S4 for this analysis.

3. Results

The search strategy found 18 papers fitting the inclusion criteria. These studies included 320 experienced drivers with driving experience ranging from 5 years to 34 years, and 318 novice drivers with driving experience ranging from 0 years to 3 years. These studies were published between 1992 and 2019, with two studies using static images of road scenes, seven studies using video clips, three driving simulator studies, five on-road studies and one study conducted on-road and in a driving simulator.

3.1. Mean fixation duration

There were ten studies that reported fixation durations as a function of experience (Chapman and Underwood, 1998a; Crundall and Underwood, 1998; Crundall et al., 1999, 2003; Huestegge et al., 2010; Jiang et al., 2012; Konstantopoulos et al., 2010; Laya, 1992; Underwood et al., 2002a; Yeung and Wong, 2015).

Two studies could not be included in the meta-analysis and therefore will be discussed narratively. One study investigated differences in experienced drivers' and novice drivers' fixation durations when displaying hazardous, high demand situations (Crundall et al., 1999) and the other investigated fixation durations around curves (Laya, 1992). These studies had a total of 28 experienced drivers and 28 novice drivers. Both studies concluded that novice drivers have significantly longer fixation durations compared to experienced drivers.

For the other eight studies, these were pooled in a meta-analysis, inputting the mean and standard deviation for both novice and experienced driver groups for each study. Although all of the included studies measured fixation durations in milliseconds, the means and standard deviations varied considerably between studies. Therefore, for this reason, and for consistency in reporting, fixation durations are firstly reported with standardised mean differences, See Fig. 2, and then with mean differences, along with an effect size (Cohen's d). Cohen's d has been calculated using the standard deviation from Chapman and Underwood (1998a) which has been chosen to be most representative due to its large and justified sample size (Higgins and Green, 2011).

Firstly, in regards to the standardised mean difference, the overall effect of driving experience on fixation durations did not produce a significant difference, $Z = 1.69$, $p = .09$, (95% CI -0.44 , 0.03), See Fig. 2.

Given that two studies included advanced drivers in the experienced group, with Crundall et al. (2003) using police pursuit drivers and Konstantopoulos et al. (2010) using driving instructors, as well as Konstantopoulos et al. (2010) and Huestegge et al. (2010) using learner drivers, a sensitivity analysis was conducted which removed these studies. When these studies were removed, there was still no overall effect of driving experience on drivers' fixation durations, $Z = 0.63$, $p = .53$ (95% CI -0.40 , 0.20), with this null result being much more evident.

Secondly, in regards to mean difference, the overall effect of driving

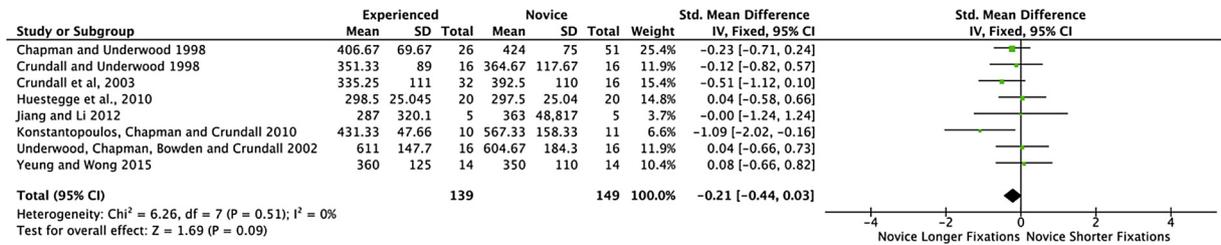


Fig. 2. A forest plot to show the standardised mean difference and overall effect of driving experience on novice and experienced drivers' fixation durations.

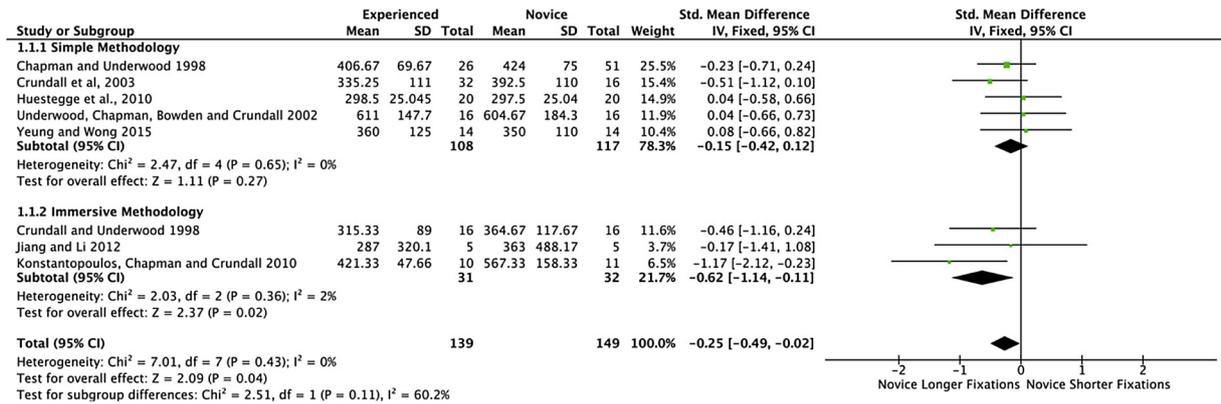


Fig. 3. A forest plot to show the subgroup analysis of how driving method (simple methodology and immersive methodology) changes the effect of experience on drivers' fixation durations.

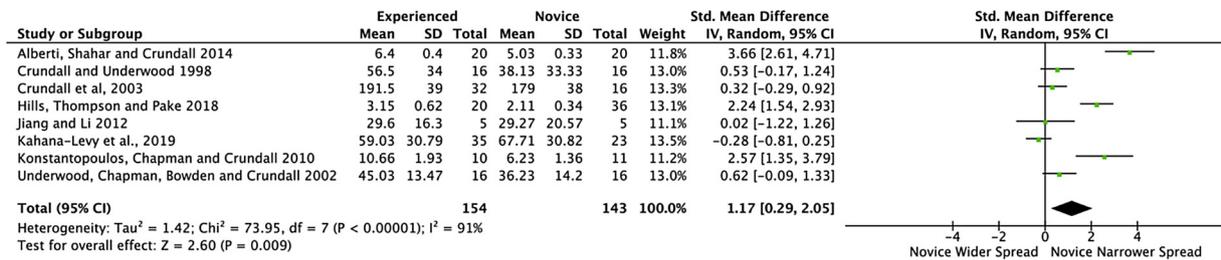


Fig. 4. A forest plot to show the overall effect of driving experience on novice and experienced drivers' mean horizontal spread of search.

experience on fixation durations did not produce a significant difference, $Z = 1.03$, $p = .30$, (95% CI -20.09, 6.24), $d = .29$.

In addition, a subgroup analysis was conducted to determine whether fixation durations vary according to the method type used. The overall effect of method type did not significantly change the effect of experience on drivers' fixation durations ($\text{Chi}^2 = 2.51$, $p = .11$). The effect of simple methodology alone was not significant ($Z = 1.11$, $p = .27$ (95% CI -0.42, 0.12)), whereas the effect of immersive methodology alone was significant ($Z = 2.37$, $p = .02$ (95% CI -1.14, -0.11)), See Fig. 3.

3.2. Horizontal spread of search

There were fifteen studies that reported horizontal spread of search as a function of experience (Alberti et al., 2014; Borowsky and Oron-Gilad, 2013; Bos et al., 2015; Chapman and Underwood, 1998a; Crundall and Underwood, 1998; Crundall et al., 1999, 2003; Hills et al., 2018; Jiang et al., 2012; Kahana-Levy et al., 2019; Konstantopoulos et al., 2010; Lehtonen et al., 2014; Mourant and Rockwell, 1972; Underwood et al., 2002a; Yeung and Wong, 2015).

Seven studies could not be included in the meta-analyses and therefore will be discussed narratively. Six of these studies, which included 121 novice drivers and 118 (incl. 10 taxi drivers) experienced drivers (Borowsky and Oron-Gilad, 2013; Bos et al., 2015; Chapman and Underwood, 1998a; Crundall et al., 1999; Lehtonen et al., 2014;

Yeung and Wong, 2015), found that there was no significant difference in novice drivers' and experienced drivers' horizontal visual search over a range of low, medium and high driving demand situations, which were conducted using a range of methods including video clips, simulators and on-road. In contrast, Mourant and Rockwell (1972) found that experienced drivers had significantly wider horizontal spread of search compared to novice drivers however, this was the only one of these seven studies to use learner drivers for the novice driver group.

For the other eight studies, these were pooled in a meta-analysis. The overall effect of driving experience on horizontal spread of search produced a significant standardised mean difference, $Z = 2.60$, $p = .009$ (95% CI 0.29, 2.05), with experienced drivers having a wider horizontal spread of search compared to novice drivers, See Fig. 4.

As before, a sensitivity analysis was conducted by removing the two studies that included an advanced experienced group and learner drivers (Crundall et al., 2003; Konstantopoulos et al., 2010). When these studies are removed, there was still an overall effect of driving experience on drivers' horizontal spread of search, $Z = 1.98$, $p < .05$ (95% CI 0.01, 2.22), however, this difference had reduced.

In addition, a subgroup analysis was conducted to determine whether horizontal spread of search varies according to the method type used. The overall effect of method type did not significantly change the effect of experience on drivers' horizontal spread of search ($\text{Chi}^2 = 1.16$, $p = .28$). The effect of simple methodology alone was not significant ($Z = 1.39$, $p = .16$ (95% CI -0.26, 1.55)), whereas the effect

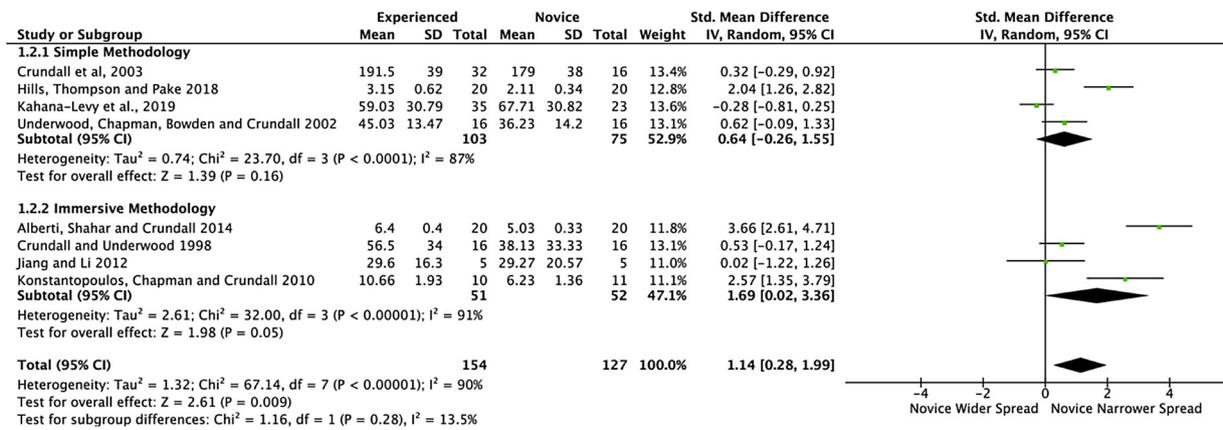


Fig. 5. A forest plot to show the subgroup analysis of how driving method (simple methodology and immersive methodology) changes the effect of experience on drivers' horizontal spread of search.

of immersive methodology alone was significant (Z = 1.98, p < .05 (95% CI 0.02, 3.36)), See Fig. 5.

3.3. Vertical spread of search

There were thirteen studies that reported vertical spread of search as a function of experience (Borowsky and Oron-Gilad, 2013; Bos et al., 2015; Chapman and Underwood, 1998a; Crundall and Underwood, 1998; Crundall et al., 1999, 2003; Jiang et al., 2012; Hills et al., 2018; Kahana-Levy et al., 2019; Konstantopoulos et al., 2010; Lehtonen et al., 2014; Mourant and Rockwell, 1972; Underwood et al., 2002a).

Five studies could not be included in the meta-analyses and therefore will be discussed narratively. Three of these studies, which included 39 novice drivers and 37 experienced drivers (Bos et al., 2015; Crundall et al., 1999; Lehtonen et al., 2014), found that there were no significant differences in novice' and experienced drivers' vertical visual search. In contrast, Mourant and Rockwell (1972) and Crundall et al. (2003) found that novice drivers had significantly wider vertical spread of search compared to experienced drivers however, these studies were the only two to use extreme driver groups in their sample in terms of police pursuit drivers and learner drivers.

For the other eight studies, these were pooled in a meta-analysis. The overall effect of driving experience on vertical spread of search did not produce a significant standardised mean difference, Z = 1.38, p = .17 (95% CI -0.68, 0.12), See Fig. 6.

Given that two of the studies included in the meta-analysis also used an advanced driver experienced group, with Borowsky and Oron-Gilad (2013) using taxi drivers and Konstantopoulos et al. (2010) using driving instructors as well as learner drivers, these studies were removed from the analysis. Again, there was no overall effect of driving experience on drivers' vertical spread of search, Z = .80, p = .42 (95% CI -0.69, 0.29).

In addition, a subgroup analysis was conducted to determine whether vertical spread of search varies according to the method type used. The overall effect of method type did not significantly change the effect

of experience on drivers' vertical spread of search (Chi² = 0.09, p = .76), with both simple methodology alone (Z = 1.21, p = .23 (95% CI -0.90, 0.21) and immersive methodology alone (Z = 0.88, p = .38 (95% CI -0.73, 0.27)) not being significant.

3.4. Number of fixations

There were seven studies that reported number of fixations as a function of experience (Borowsky and Oron-Gilad, 2013; Bos et al., 2015; Crundall and Underwood, 1998; Huestegge et al., 2010; Kahana-Levy et al., 2019; Konstantopoulos et al., 2010; Underwood et al., 2003).

Two studies could not be included in the meta-analyses and therefore will be discussed narratively. These two studies, which include 27 novice drivers and 35 experienced drivers (incl. 10 taxi drivers) (Borowsky and Oron-Gilad, 2013; Bos et al., 2015), found that there was no significant difference between the number of fixations made by experienced drivers and novice drivers.

For the other five studies, these were pooled in a meta-analysis. The overall effect of driving experience for number of fixations did not produce a significant standardised mean difference, Z = 1.10, p = .27 (95% CI -0.13, 0.46).

A sensitivity analysis was conducted by removing the two studies that included learner drivers, with one of these studies also including driving instructors (Huestegge et al., 2010; Konstantopoulos et al., 2010). By removing these studies, the heterogeneity involved in this meta-analysis was reduced from 13% to 0%. The removal of these studies did not change the overall effect dramatically, still failing to produce a significant difference, Z = .96, p = .34 (95% CI -0.18, 0.54).

In addition, a subgroup analysis was conducted to determine whether the number of fixations varies according to the method type used. The overall effect of the method type used in the study did not significantly change the effect of experience on drivers' number of fixations (Chi² = .21, p = .64), with the effect of simple methodology alone

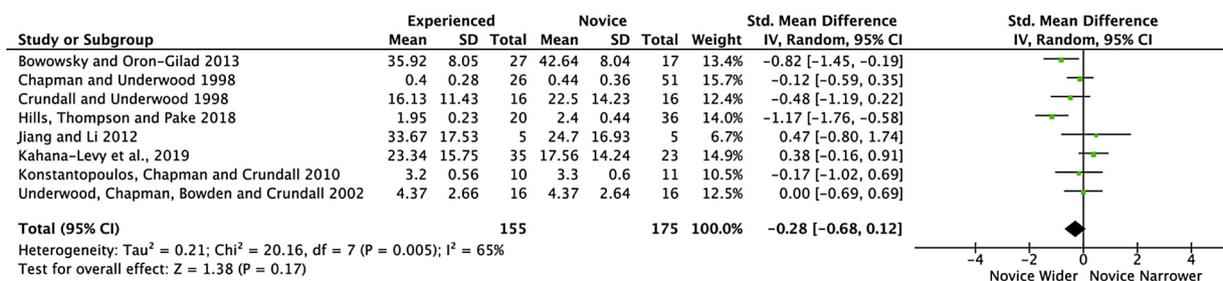


Fig. 6. A forest plot to show the overall effect of driving experience on novice' and experienced drivers' vertical spread of search.

($Z = .32$, $p = .75$ (95% CI -0.49 , 0.69) and immersive methodology alone ($Z = .91$, $p = .36$ (95% CI -0.35 , 0.96)) not being significant.

4. Discussion

4.1. Summary of results

Despite claims for the past 40 years that novice drivers' visual search differs from experienced drivers' visual search on the road, the current findings suggest that these differences are not so apparent when all available studies are pooled together. While it was clear that there was a difference in drivers' horizontal spread of search, with novice drivers having a narrower horizontal spread of search compared to experienced drivers, there were no reliable differences found in fixation durations, vertical spread of search and number of fixations. While horizontal spread of search continues to support the general conclusions from previous literature, there are some factors that need to be considered when identifying the differences between novice drivers' and experienced drivers' eye movements such as the experience level of the two groups of drivers, and the effect of method type.

4.2. Differences in novice drivers' and experienced drivers' visual search

The only measure to show a clear overall difference in the meta-analysis between novice and experienced drivers was horizontal spread of search, with novice drivers displaying a narrower spread of search compared to experienced drivers. This difference has been widely interpreted as novice drivers having limited experience in scanning and anticipating the location of potential hazards in the peripheral (e.g. Mourant and Rockwell, 1972; Kahana-Levy et al., 2019).

In addition, the sensitivity analysis which temporarily removed studies that investigated extreme experience groups showed a reduction in the difference between novice and experienced drivers, suggesting that the inclusion of these drivers may have accounted for a substantial proportion of the overall effect. The inclusion of these groups may have also accounted for the differences in the results of the studies described narratively (Mourant and Rockwell, 1972). However, without the inclusion of these groups, the difference between novice and experienced drivers still remained.

When the studies were sub-grouped by method type, this factor was not seen to change the effect of experience on drivers' horizontal spread of search, with both simple and immersive sub-groups displaying a trend towards novice drivers having a narrower spread of search than experienced drivers. However, when focussing on the subgroups individually, the studies conducted in an immersive environment produced a significant difference between novice' and experienced drivers' horizontal spread of search whereas, simple methodology did not. This finding is supported by previous research which indicates that only experienced drivers make use of a more immersive, wider field of view to detect oncoming hazards, with novice drivers failing to look for potential hazards in the peripheral (Alberti et al., 2014).

4.3. Absence of differences in novice drivers' and experienced drivers' visual search

In regards to fixation durations, the pooling of relevant studies in the meta-analysis showed no overall difference between novice drivers and experienced drivers, refuting the widely used claim that novice drivers have generally longer fixation durations over the visual scene compared to experienced drivers (e.g. Rayner, 1998; Green, 2007). Although there were no overall differences, this result should be interpreted with caution as it is not as conclusive as other measures. The overall effect size (Cohen's d of 0.29), which can be calculated for this measure given the compatibility of units between studies, can be seen to be between small and medium in Cohen's terms (Cohen, 1988).

When temporarily removing the studies that included drivers with

extreme experience, the absence of an overall difference became much more pronounced for fixation durations compared to the removal of these studies for all other measures. This suggests that extreme experience groups may be driving the tendency towards a difference between novice and experienced drivers' fixation durations, which could have important practical implications for interventions.

In addition, while the subgrouping of studies by method type did not produce a significant overall difference, the studies conducted in an immersive environment produced a significant difference between novice and experienced drivers' fixation durations, whereas simple methodology did not. These findings suggest that when further studies are conducted that involve driving on real roads it is possible that a reliable difference in fixation durations between novice and experienced drivers may yet emerge. The absence of an overall difference in the current meta-analysis may be driven by the majority of studies being conducted using simple methodology, so it remains possible that drivers' fixation durations in immersive driving situations may still be relevant in predicting and reducing accident involvement for novice drivers on real roads.

In regards to vertical spread of search, the pooling of all relevant studies revealed no overall difference between novice and experienced drivers, refuting previous research which has found that newly qualified drivers favour vertical search due to different information needs, i.e. helping maintain lane position (Land and Horwood, 1995). This lack of difference between the two groups is perhaps understandable, as vertical spread of search is arguably less important compared to horizontal search in a driver's ability to detect hazards, and therefore neither experienced, nor novice drivers are searching wider than deemed necessary (Chapman and Underwood, 1998a). In addition, as there was no effect of method type, these findings suggest that the measure of vertical spread of search in all contexts is not sensitive enough to demonstrate differences between the two groups, and therefore is not a reliable measure to include in visual search training interventions.

The sensitivity analysis, which removed the studies that included learner drivers and advanced driver groups, was seen to further confirm that there were no differences between the two groups. The most influential study which indicated differences in vertical spread of search (Mourant and Rockwell, 1972) has also been previously criticised for the minimal experience their learner drivers had on the road, and therefore an alternative interpretation for the increase in vertical spread of search demonstrated by novice drivers could be due to difficulties in controlling the vehicle. This finding suggests that previous reports of differences in novice and experienced drivers' vertical spread of search may not be representative of typical changes in visual search over the first years of unsupervised driving.

Finally, it was found that there was no overall difference in novice drivers' and experienced drivers' number of fixations over the driving scene. This pooling of data is in contrast to previous claims which report that novice drivers make fewer fixations compared to experienced drivers in order to limit the amount of visual information being processed (Crundall and Underwood, 1998). When these data were sub-grouped by method type, this was not seen to change the effect of experience. This increases the reliability of this finding, with this absence of a difference not being sensitive to the immersiveness of the environment.

In light of the 'where' and 'when' pathways (Findlay and Walker, 1999), the absence of an overall effect for number of fixations is compatible with the absence of an effect for fixation durations. That said, this systematic review only focuses on the number of fixations drivers make generally over the visual scene while completing a task, and not on specific areas of interest such as the number of fixations drivers make on their external mirrors or at wide eccentricities, due to the limited number of studies that have directly investigated this (Underwood et al., 2002b; Konstantopoulos et al., 2010). It is these areas of interest that may be related to driving experience, particularly

as novice and experienced drivers' horizontal spread of search significantly differed over the visual scene.

4.4. Implications of results

The findings from the current systematic review have a number of implications for road user safety. Firstly, these findings help to understand the high accident liability of novice drivers, by highlighting the potential problems in their cognitive processing, reflected by their eye movements. Secondly, these visual search differences help with the development of interventions, with results suggesting that horizontal spread of search should be the immediate focus when developing training interventions for novice drivers. For example, by encouraging novice drivers to have a wider visual search in order to scan for potential hazards, as well as repeated exposure to hazards that could develop in the periphery, this could help improve novice drivers' knowledge and understanding in such situations (Chapman et al., 2002).

While previous authors have cautioned against a wholesale encouragement of a broader spread of search, in case this results in incomplete processing of the objects or locations being currently fixated (Chapman et al., 2002), the current results suggest that differences in processing times between novice drivers and experienced ones may be relatively small compared to differences in horizontal spread of search. This does highlight the importance of top down influences on the "where" pathway and suggests that interventions can safely encourage new drivers to devote their search to a wider range of horizontal locations in the visual scene.

In contrast, visual training interventions should have less emphasis on improving vertical spread of search, for example increasing drivers' ability to look further down the road. This would require the changing of current practical training interventions, such as the Road Craft manual for police drivers (Coyne et al., 2007) that indicates that drivers should 'increase the length and breadth of their vision' and 'The Roadcraft Education Strategy' to educate non-drivers or learner drivers, encouraging 'forward observation and peripheral vision awareness' (The Roadcraft Model, 2018).

Finally, these findings also highlight some factors that should be carefully considered when conducting future research studies. Both the experience and training of the recruited driver groups should be an important consideration when predicting and interpreting results based on previous literature. In addition, studies which use simple methodology to compare visual search strategies, particularly when measuring drivers' fixation durations and horizontal spread of search, should be mindful of the fact that the absence of differences found between groups may not be representative of drivers' behaviour in more immersive driving environments and on real roads.

4.5. Limitations of the studies

When critically appraising the included studies, this process highlighted limitations in the field that could be addressed when conducting further research. Firstly, in terms of driver recruitment there are often problems generalising from academic research using student samples on to the broader population. In this case, although we have included some studies involving participants from a student population, many of the studies have recruited novice drivers direct from Test Centres, and experienced drivers from Newspaper advertisements. Secondly, the majority of included studies did not report any drop outs during the study. This is particularly surprising for driving simulator studies, as there is a high likelihood that some participants would have dropped out due to simulator sickness (Brooks et al., 2010). Future research studies should clearly report participant dropout rates in order to give an accurate representation of the included sample.

In addition, all of the studies, with the exception of Chapman and Underwood (1998a), gave no justification of sample size. This lack of

justification allows for underpowered experiments to be conducted, with the danger that null results from these relatively small studies cannot be published. The associated implication is publication bias, with the literature being over-represented by studies with positive results. Although publication bias is a potential source of bias, formal tests could not be performed due to the number of studies included for each outcome measure. The guidelines for the use of funnel plots and asymmetry tests vary from at least 10 studies to an ideal number of 75 studies for high power (Higgins & Green; Begg and Mazumdar, 1994). However, to help minimise publication bias, the current search strategy had no restrictions on publication status by including research in progress databases. In addition, this is an area in which studies with relatively small sample sizes may still be published in some form given the effort required to obtain the data. In fact, the study with the smallest sample size (Jiang et al., 2012) is one with the smallest effect sizes for any of the key variables. Moreover, given that the only study to justify the sample size was not seen as an outlier in any of the outcome measures, there is no direct evidence of effect-size inflation in the current data.

Finally, it was noted at the start of the review that although there are many studies investigating this topic, it is difficult to compare these due to potential forms of heterogeneity in terms of design, outcomes measures and participants. In terms of the design, the demands of the driving task varied dramatically between the included studies, with the use of many different road types (e.g. rural, dual carriageway and curved roads) which could not be operationally defined in order to pool the studies. A second form of potential heterogeneity was the measurement of outcomes. Even though fixation duration was the only outcome to be measured in consistent units, the values still varied substantially across studies. It is possible that these differences are due to factors such as authors adopting different eye tracker dispersion algorithms, with this criterion not always being documented in research outputs. While these forms of heterogeneity cannot be avoided in such a review, this stresses the importance of using standardised mean differences and random effects models to account for this. In regards to participants, it has previously been highlighted that there is no consistent definition of a novice or experienced driver in the literature, and therefore a distinction was made based on UK crash statistics (Clarke et al., 2006). However, the sensitivity analyses did confirm that none of our overall conclusions would have been different even if a more restrictive definition of novice and experienced drivers had been adopted.

4.6. Conclusion

In summary, the pooling of studies in this systematic review provides reliable conclusions regarding the difference between novice drivers' and experienced drivers' visual search, with novice drivers displaying a narrower spread of horizontal search compared to more experienced drivers, suggesting that novice drivers do not anticipate and scan for potential hazards to either side of them. In contrast, no reliable experience differences were found for fixation durations, vertical spread of search and number of fixations. A key implication for the development of training interventions is that novice drivers need to develop a broader horizontal spread of visual search but do not necessarily need to learn to "look further down the road". Limitations in novice drivers' fixation durations, and to some extent horizontal search, are most notable for learner drivers and in immersive testing environments, therefore we recommend that this should be the focus for future research, training, and evaluation.

Declaration of Competing Interest

None.

Acknowledgements

This work was supported by the Economic and Social Research Council [grant number ES/J500100/1], an ESRC post-graduate studentship to Chloe Robbins. We would like to thank Adam Sellers who conducted an independent search of the literature to ensure that all relevant studies had been identified, and served as the second reviewer for the relevance of potentially included studies.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.aap.2019.105266>.

References

- Alberti, C.F., Gamberini, L., Spagnoli, A., Varotto, D., Semenzato, L., 2012. Using an eye-tracker to assess the effectiveness of a three-dimensional riding simulator in increasing hazard perception. *Cyberpsychol. Behav. Soc. Netw.* 15 (5), 274–276.
- *Alberti, C.F., Shahar, A., Crundall, D., 2014. Are experienced drivers more likely than novice drivers to benefit from driving simulations with a wide field of view? *Transp. Res. Part F Traffic Psychol. Behav.* 27, 124–132.
- Antman, E.M., Lau, J., Kupelnick, B., Mosteller, F., Chalmers, T.C., 1992. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *Jama* 268 (2), 240–248.
- Begg, C.B., Mazumdar, M., 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1088–1101.
- Bingham, C.R., Shope, J.T., 2004. Adolescent problem behavior and problem driving in young adulthood. *J. Adolesc. Res.* 19 (2), 205–223.
- *Borowsky, A., Oron-Gilad, T., 2013. Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accid. Anal. Prev.* 59, 548–565.
- Bos, A.J., Ruscio, D., Cassavaugh, N.D., Lach, J., Gunaratne, P., Backs, R.W., 2015. Comparison of novice and experienced drivers using the SEEV model to predict attention allocation at intersections during simulated driving. *Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment*.
- Bremmer, F., Kubischik, M., Hoffmann, K., Krekelberg, B., 2009. Neural dynamics of saccadic suppression. *J. Neurosci.* 40, 12374–12383.
- Brooks, J.O., Goodenough, R.R., Crisler, M.C., Klein, N.D., Alley, R.L., Koon, B.L., et al., 2010. Simulator sickness during driving simulation studies. *Accid. Anal. Prev.* 42 (3), 788–796.
- Cantwell, S., Isler, R., Starkey, N., 2013. The effects of road commentary training on novice drivers' visual search behaviour: a preliminary investigation. Paper Presented at the Australasian Road Safety Research Policing Education Conference.
- 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. *Saf. Sci.* 82, 144–154.
- *Chapman, P.R., Underwood, G., 1998a. Visual search of driving situations: danger and experience. *Perception* 27, 951–964.
- Chapman, P.R., Underwood, G., 1998b. Visual search of dynamic scenes: event types and the role of experience in viewing driving situations. In: Underwood, G. (Ed.), *Eye Guidance in Reading and Scene Perception*. Elsevier, Oxford, pp. 371–396.
- Chapman, P., Underwood, G., Roberts, K., 2002. Visual search patterns in trained and untrained novice drivers. *Transp. Res. Part F Traffic Psychol. Behav.* 5 (2), 157–167.
- Chisholm, S.L., Caird, J.K., Lockhart, J., 2008. The effects of practice with MP3 players on driving performance. *Accid. Anal. Prev.* 40 (2), 704–713.
- Clarke, D.D., Ward, P., Bartle, C., Truman, W., 2006. Young driver accidents in the UK: the influence of age, experience, and time of day. *Accid. Anal. Prev.* 38 (5), 871–878.
- Cohen, A.S., 1981. Car drivers' pattern of eye fixations on the road and in the laboratory. *Percept. Mot. Skills* 52 (2), 515–522.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioural Sciences*, second edition. Lawrence Erlbaum, Hillsdale, New Jersey.
- Cooper, H., 1998. *Cooper, Harris, Synthesizing Research: A Guide for Literature Reviews*. Sage, Thousand Oaks, CA.
- Coyne, P., Foundation, P., Mares, P., Macdonald, B., 2007. *Roadcraft: The Police Driver's Handbook*. The Stationery Office, London.
- Crundall, D., Andrews, B., Van Loon, E., Chapman, P., 2010. Commentary training improves responsiveness to hazards in a driving simulator. *Accid. Anal. Prev.* 42 (6), 2117–2124.
- *Crundall, D., Chapman, P., Phelps, N., Underwood, G., 2003. Eye movements and hazard perception in police pursuit and emergency response driving. *J. Exp. Psychol. Appl.* 9 (3), 163.
- Crundall, D., Chapman, P., Trawley, S., Collins, L., Van Loon, E., Andrews, B., Underwood, G., 2012. Some hazards are more attractive than others: drivers of varying experience respond differently to different types of hazard. *Accid. Anal. Prev.* 45, 600–609.
- *Crundall, D.E., Underwood, G., 1998. Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics* 41 (4), 448–458.
- Crundall, D., Underwood, G., Chapman, P., 1999. Driving experience and the functional field of view. *Perception* 28 (9), 1075–1087.
- Crundall, D., Underwood, G., Chapman, P., 2002. Attending to the peripheral world while driving. *Appl. Cognit. Psychol.* 16 (4), 459–475.
- Deery, H.A., 1999. Hazard and risk perception among young novice drivers. *J. Safety Res.* 30 (4), 225–236.
- DfT, 2015. *Facts on Young Car Drivers: June 2015*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/448039/young-car-drivers-2013-data.pdf.
- DfT, 2018. *Contributory Factors for Reported Road Accidents*. Retrieved from: <https://www.gov.uk/government/statistical-data-sets/ras50-contributory-factors>.
- Di Stasi, L.L., Contreras, D., Cándido, A., Cañas, J.J., Catena, A., 2011. Behavioral and eye-movement measures to track improvements in driving skills of vulnerable road users: first-time motorcycle riders. *Transp. Res. Part F Traffic Psychol. Behav.* 14 (1), 26–35.
- Drews, F.A., Pasupathi, M., Strayer, D.L., 2008. Passenger and cell phone conversations in simulated driving. *J. Exp. Psychol. Appl.* 14 (4), 392.
- Findlay, J.M., Walker, R., 1999. A model of saccade generation based on parallel processing and competitive inhibition. *Behav. Brain Sci.* 22 (4), 661–674.
- Fisher, D.L., Caird, J., Horrey, W., Trick, L. (Eds.), 2016. *Handbook of Teen and Novice Drivers: Research, Practice, Policy, and Directions*. CRC Press.
- Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T.J., et al., 2011. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *J. Clin. Epidemiol.* 64 (11), 1187–1197.
- Glass, G.V., 1976. Primary, secondary, and meta-analysis of research. *Educ. Res.* 5 (10), 3–8.
- Green, P., 2007. Where do drivers look while driving (and for how long)? In: Dewar, R., Olson, P. (Eds.), *Human Factors in Traffic Safety*, 2nd eds. Lawyers and Judges Publishing, Tucson, AZ, pp. 57–86.
- Haigney, D., 1995. *STATS19. INROADS 17* (2).
- Hartling, L., Wiebe, N., Russell, K.F., Petruk, J., Spinola, C., Klassen, T.P., 2004. Graduated driver licensing for reducing motor vehicle crashes among young drivers. *Cochrane Database Syst. Rev.* 2.
- Higgins, J.P., Green, S. (Eds.), 2011. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration Retrieved from: www.handbook.cochrane.org.
- *Hills, P.J., Thompson, C., Pake, J.M., 2018. Detrimental effects of carryover of eye movement behaviour on hazard perception accuracy: effects of driver experience, difficulty of task, and hazardousness of road. *Transp. Res. Part F Traffic Psychol. Behav.* 58, 906–916.
- Horswill, M.S., Taylor, K., Newnam, S., Wetton, M., Hill, A., 2013. Even highly experienced drivers benefit from a brief hazard perception training intervention. *Accid. Anal. Prev.* 52, 100–110.
- *Huestegge, L., Skottke, E.M., Anders, S., Müsseler, J., Debus, G., 2010. The development of hazard perception: dissociation of visual orientation and hazard processing. *Transp. Res. Part F Traffic Psychol. Behav.* 13 (1), 1–8.
- Isler, R.B., Starkey, N.J., Sheppard, P., 2011. Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. *Accid. Anal. Prev.* 43 (5), 1818–1827.
- *Jiang, J., Lu, J., Li, X.Q., 2012. Eye movement of Chinese experienced drivers and novice drivers. *Transportation Research Board 91st Annual Meeting*, (No. 12-2836).
- *Kahana-Levy, N., Shavitzky-Golkin, S., Borowsky, A., Vakil, E., 2019. The effects of repetitive presentation of specific hazards on eye movements in hazard perception training, of experienced and young-inexperienced drivers. *Accid. Anal. Prev.* 122, 255–267.
- *Konstantopoulos, P., Chapman, P., Crundall, D., 2010. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accid. Anal. Prev.* 42 (3), 827–834.
- Konstantopoulos, P., Chapman, P., Crundall, D., 2012. Exploring the ability to identify visual search differences when observing drivers' eye movements. *Transp. Res. Part F Traffic Psychol. Behav.* 15 (3), 378–386.
- Land, M., Horwood, J., 1995. Which parts of the road guide steering? *Nature* 377 (6547), 339–340.
- Law, M., Stewart, D., Pollock, N., Letts, L., Bosch, J., Westmorland, M., 1998. *Critical Review form-Quantitative Studies*. McMaster University: Occupational Therapy Evidence-Based Practice Research Group.
- Laya, O., 1992. Eye movements in actual and simulated curve negotiation tasks. *IATSS Res.* 16 (1), 15–26.
- Lehtonen, E., Lappi, O., Koirikivi, I., Summala, H., 2014. Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accid. Anal. Prev.* 70, 195–208.
- Mayhew, D.R., Simpson, H.M., Pak, A., 2003. Changes in collision rates among novice drivers during the first months of driving. *Accid. Anal. Prev.* 35, 683–691.
- McCartt, A.T., Mayhew, D.R., Braitman, K.A., Ferguson, S.A., Simpson, H.M., 2009. Effects of age and experience on young driver crashes: review of recent literature. *Traffic Inj. Prev.* 10 (3), 209–219.
- McCartt, A.T., Shabanova, V.I., Leaf, W.A., 2003. Driving experience, crashes, and traffic citations of teenage beginning drivers. *Accid. Anal. Prev.* 35, 311–320.
- Miller, J., Stacey, M., 2013. *The Driving Instructor's Handbook*. Kogan Page Publishers.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151 (4), 264–269.
- Mourant, R.R., Rockwell, T.H., 1972. Strategies of visual search by novice and experienced drivers. *Hum. Factors* 14 (4), 325–335.
- National Highway Traffic Safety Administration, 2016. *Traffic Safety Facts: Young Drivers*. Retrieved from: <https://crashstats.nhtsa.dot.gov/Api/Public/Publication/812498>.
- Peck, R.C., 1993. The identification of multiple accident correlates in high risk drivers

- with specific emphasis on the role of age, experience, and prior traffic violation frequency. *Alcohol Drugs Driving* 9, 145–166.
- Polanin R., J., Pigott D., T., 2015. The use of meta-analytic statistical significance testing. *Research synthesis methods* 6 (1), 63–73.
- Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124 (3), 372.
- Renge, K., 1980. The effect of driving experience on driver's visual attention: an analysis of objects looked at, using the verbal report method. *IATSS Res.* 4 (1), 95–106.
- Ryan, R., 2013. Cochrane Consumers and Communication Review Group: Data Synthesis and Analysis. Cochrane Consumers and Communication Review Group.
- The Roadcraft Model, 2018. Roadcraft Driver Education Specialists [PDF file]. Retrieved from: <https://www.aph.gov.au/DocumentStore.ashx?id=67352f51-8e36-4773-829d-d98e134ff97c>.
- Sabey, B.E., Staughton, G.C., 1975. Interacting roles of road environment vehicle and road user in accidents. *Proceedings of the Fifth International Conference of the Association for Accident and Traffic Medicine* 1–17.
- Scammacca, N., Roberts, G., Stuebing, K.K., 2014. Meta-analysis with complex research designs: dealing with dependence from multiple measures and multiple group comparisons. *Rev. Educ. Res.* 84 (3), 328–364.
- Schünemann, H., Brożek, J., Guyatt, G., Oxman, A., 2013. Handbook for Grading the Quality of Evidence and the Strength of Recommendations Using the GRADE Approach. GRADE Working Group.
- Underwood, G., 2007. Visual attention and the transition from novice to advanced driver. *Ergonomics* 50 (8), 1235–1249.
- *Underwood, G., Chapman, P., Bowden, K., Crundall, D., 2002a. Visual search while driving: skill and awareness during inspection of the scene. *Transp. Res. Part F Traffic Psychol. Behav.* 5 (2), 87–97.
- *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.
- Underwood, G., Crundall, D., Chapman, P., 2002b. Selective searching while driving: the role of experience in hazard detection and general surveillance. *Ergonomics* 45 (1), 1–12.
- World Health Organization, 2018. Global Status Report of Road Safety 2018. World Health Organization.
- *Yeung, J.S., Wong, Y.D., 2015. Effects of driver age and experience in abrupt-onset hazards. *Accid. Anal. Prev.* 78, 110–117.