The effectiveness of advanced driver training

July 2021

V Beanland, I Huemmer

Waka Kotahi NZ Transport Agency research report 677
Contracted research organisation – University of Otago

ISBN 978-1-99-004421-2 (electronic)
ISSN 1173-3764 (electronic)

Waka Kotahi NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
NZTAresearch@nzta.govt.nz
www.nzta.govt.nz

Beanland, V and I Huemmer (2021) The effectiveness of advanced driver training. *Waka Kotahi NZ Transport Agency research report 677*.

The University of Otago was contracted by Waka Kotahi NZ Transport Agency in 2020 to carry out this research.

Keywords: driver training, driver education, graduated driver licensing, novice drivers, teen drivers, young drivers

An important note for the reader

Waka Kotahi NZ Transport Agency is a Crown entity established under the Land Transport Management Act 2003. The objective of Waka Kotahi is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, Waka Kotahi funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research and should not be regarded as being the opinion or responsibility of Waka Kotahi. The material contained in the reports should not be construed in any way as policy adopted by Waka Kotahi or indeed any agency of the New Zealand Government. The reports may, however, be used by New Zealand Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation, Waka Kotahi and agents involved in their preparation and publication do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents of the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or other expert advice.

Erratum

21 February 2022

Page 60, section 5.1, first paragraph, reference corrected: from Schiff 2019 to Begg and Brookland 2015.

Acknowledgements

Thank you to the following people for their input and assistance to this project and the preparation of this report:

- Steering Group chair, Malcolm Menzies from Waka Kotahi NZ Transport Agency
- Steering Group members Alex McMinn from Waka Kotahi NZ Transport Agency, and Marilyn Stephens and Matthew Stone from the Ministry of Transport
- Professor Samuel Charlton from the University of Waikato and Dr Lyndel Bates from Griffith University, who conducted peer review of the report.

Abbreviations and acronyms

AAA	American Automobile Association
AAC	L’Apprentissage anticipé de la conduite (France)
AAHPT	Act and Anticipate Hazard Perception Training
ACCEL	Accelerated Curriculum to Create Effective Learning
ACT	Anticipate–Control–Terminate
EDTS	Engaged Driver Training System
FOCAL	Forward Concentration and Attention Learning
GDLS	Graduated driver licensing system
GMOST	Game-based multi-user online simulated training
PDL	Pre-Driver Licensing (programme in DeKalb County, Georgia, USA)
POP	Provisional operator’s permit (Nebraska, USA)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RACC	Real Automobil Club de Cataluña (Spain)
RAPT	Risk awareness and perception training
SPC	Safe Performance Curriculum (programme in DeKalb County, Georgia, USA)
STRAP	Secondary Task Regulatory and Anticipatory Program
TRID	Transportation Research Information Database
ZED	Zero Errors Driving

Contents

Executive summary

Advanced driver training courses are designed to teach skills that can improve vehicle handling and hazard perception, which could theoretically help reduce the likelihood of being involved in a crash. For this reason, in New Zealand, graduates of approved advanced driver training courses are currently permitted to apply for their full licence earlier than drivers who have not completed an approved course. New Zealand has a graduated driver licensing system (GDLS) in which drivers must first obtain a conditional learner licence and hold it for at least 6 months before they can apply for a restricted licence. Drivers under 25 years of age must hold their restricted licence for at least 18 months, and drivers aged 25 and over must hold their restricted licence for at least 6 months. However, these minimum durations can be reduced to 12 and 3 months, respectively, after completing an approved driver training course.

A recent evaluation of the New Zealand GDLS (Schiff 2019) indicated that overall crash rates were not significantly different when comparing trained and untrained drivers. However, crash rates were significantly lower for a subset of the trained drivers who had not accepted the time discount. This suggests that among drivers who did accept the time discount, any potential safety benefits of driver training were counteracted by earlier licensure.

The purpose of the research reported here was to investigate the approaches used currently across the world in advanced driver training and review the effects that different training approaches have on drivers' skills, safety, behaviour and attitudes. It did not set out to evaluate the programmes delivered within the New Zealand Advanced Driver Training System (see Appendix). This project was conducted over the period November 2020 to March 2021.

Method

A systematic review of the driver training literature was undertaken. Three scholarly databases were searched (PsycINFO, Scopus, and the Transport Research Information Database) using 'driver training' and 'driver education' as keywords. These searches returned 4,472 items, of which 824 were discarded as duplicates. A further 3,396 items were deemed out of scope based on title and abstract review, and for 39 items the full text was unavailable. This left 213 items for which the full text was reviewed. The full text review identified 7 additional relevant publications, but 138 items were deemed out of scope or redundant, resulting in a final set of 82 publications included in the systematic review. Papers were grouped into broad categories based on the type of training evaluated.

Training incorporated within the licensing process

Several jurisdictions incorporate either compulsory or optional training within the licensing process, usually at the pre-licence or learner stage. In some cases, the training programme represents one of two alternative pathways to licensure, whereas in other cases completing training makes the individual eligible for earlier licensure or a time discount on their licence.

Compulsory training has been evaluated in Finland, Denmark and Canada, with different results in each country; this is because the nature of the training and the consequences of making it compulsory differed in each country. The Danish evaluation found a significant reduction in multi-vehicle and manoeuvring crashes after the introduction of a new, structured training programme that emphasised hazard perception. However, the Canadian evaluation found a significant increase in crashes among young females after driver training was made compulsory in Quebec, because the new policy encouraged earlier licensing. Finally, the Finnish evaluation found no significant change in slippery crash rates following the introduction of compulsory skid training.

Time discount policies have been evaluated in Canada and in New Zealand (Schiff 2019). These evaluations mostly found negative effects of a time discount (ie, increased crashes and violations) with some analyses indicating no significant impact. No studies found a positive safety impact of time discounts. Similarly, a study in the Netherlands found that an intensive driver training programme, which effectively allowed learner drivers to get their licence sooner, was associated with significantly more incidents after licensure.

Finally, studies in USA have evaluated driver training as an alternative or modified pathway for licensure. In Nebraska, completing a driver training course removes the requirement to log 50 hours of supervised driving, and in Oregon, completing an approved training course reduces the supervised practice requirement from 100 to 50 hours. Evaluations of both policies found significantly fewer crashes and violations among trained drivers. Notably, these courses do not reduce licence duration, and it is unclear how much practice the learner drivers actually obtained. That is, although the training courses ostensibly replaced 50 hours of supervised practice, drivers may have done training in addition to practice.

Programmes incorporating on-road training

Over the past 40 years there have been multiple evaluations of driver training courses that incorporate both practical lessons and classroom- or discussion-based sessions, some of which use large samples (thousands of drivers). Most of the programmes were relatively short courses (1–2 days) aimed at licensed drivers, but some were longer courses (20–50 hours training) designed for learner drivers.

Evaluations of programmes incorporating on-road training have yielded inconsistent results, with most studies showing that training either has no significant impact on safety, or a negative safety impact. These patterns have occurred both across studies that randomly assign drivers to receive training, and studies in which drivers volunteer or self-select to undertake training. A key issue is that on-road training emphasises procedural vehicle handling skills; even when these skills are improved, it does not necessarily translate into reduced crash rates.

Simulator training

Several studies have evaluated simulator training for drivers at different stages. An advantage of simulator training is that learners can be safely and efficiently exposed to a diverse range of situations, including critical situations that could result in a crash and scenarios that would be relatively uncommon on real roads. Scenarios can be designed to progressively increase in difficulty and can be tailored to the individual so that they receive more training on skills that require more development.

Evaluations have found that simulator training is effective at developing procedural skills in learner drivers, and can improve later on-road safety, especially if the simulator uses a wide field of view (≥ 135°); however, not all studies find a significant difference between trained and untrained drivers. For licensed drivers, simulator training appears to be most beneficial when it is tailored to provide feedback on the individual's specific weaknesses. A major limitation of simulator training studies is that most use relatively small sample sizes, and the samples are not representative, which limits the generalisability of results.

Hazard perception training

Research on hazard perception training has increased dramatically since the 1990s, inspired by studies indicating that traditional (ie, on-road) driver training did not significantly improve safety. There are many different types of hazard perception programmes; some address hazard anticipation and response in a broad sense, and some address specific aspects of hazard perception (eg, rural road hazards). Most hazard perception programmes are computer-based, involving watching videos or completing simulated exercises containing key hazard events, and some programmes have been delivered online. Training programmes are typically very short, usually less than 1 hour with some lasting only 10 minutes.

Approximately half the evaluations identified in the current systematic review involved hazard perception training. However, nearly all these evaluations were small-scale experimental studies, some with as few as 10 participants per group, and most only examined short-term training effects (ie, immediately post-training or 1 week later). Most evaluations find that hazard perception training yields positive effects, measured as improved hazard perception and more appropriate glance behaviour (eg, more glances and earlier glances at potential hazard locations, fewer long off-road glances). However, few studies have assessed whether this translates to reduced crash rates during real driving, and studies that have more directly assessed driver safety have not been as conclusive or positive as those that examined only hazard perception. Some studies have indicated a ‘dose-response’ effect; that is, that refresher courses help maintain the effects of hazard perception training. More research is needed using larger samples to confirm the conditions required for hazard perception training to significantly improve on-road safety.

Other training

A few training programmes did not fit into any of the previously mentioned categories, but are worth mentioning. These programmes are diverse, but most address psychosocial aspects of driving, such as resilience (eg, resisting peer pressure to drink drive), peer relationships (eg, fostering safer communication between young drivers and peer passengers) and cognitive factors that influence speeding. Most of these programmes showed positive effects; however, given the limited research on these programmes, all findings should be considered as preliminary (especially for programmes that have only been subject to smaller-scale experimental evaluations).

Conclusions and recommendations

Although there is a large body of research evaluating driver training, most evaluations have significant limitations and problems, so it is difficult to draw definitive conclusions about the effectiveness of driver training. However, several evaluations have suggested that time discounts do more harm than good (eg, by increasing violation rates and/or crash risk), and therefore we recommend that the current restricted phase ‘time discount’ following completion of an approved driver training course should be removed. If this recommendation is followed, the impact of the change should be monitored to assess its effects on both safety and broader social issues such as access to employment (as young drivers would have to wait longer to obtain a full unrestricted licence).

Given the significant limitations noted in relation to the quality of both training programmes and evaluations, we make several recommendations to improve future driver training programmes and evaluations.

Driver training programmes should have a clearly articulated aim (ie, a specific issue or skill set they are intended to address). They should be experiential rather than based solely on observation, instruction or theory, and should be designed so that they highlight drivers’ limitations rather than being framed as ‘skill improvement’. Training should involve distributed practice over an extended time period, rather than intensive practice over a short period. Crucially, training programmes should not enable early licensure.

Training programmes should take into account individual differences in personality, driving style and motivations for driving, as previous research has indicated that training programmes are only effective for ‘careful’ drivers. This may require development of entirely new programmes that are tailored to drivers who have intrinsically low safety motivation.

The content of training programmes should be recorded in detail to ensure that programmes can be delivered consistently by different training providers. Finally, we recommend that training programmes be appropriately evaluated. All evaluations should include an appropriate comparison group and should include measures to assess the training's effectiveness on both target skills (eg, vehicle handling, hazard perception) and on-road safety (eg, crashes, violations).

Abstract

Advanced driver training courses are designed to teach skills that can improve vehicle handling and hazard perception, which could theoretically reduce crash risk. In the current New Zealand graduated driver licensing system, Class 1 (car) licence holders can obtain their full licence earlier if they complete an approved driver training course. However, a recent evaluation revealed that crash rates were not significantly different between untrained drivers and trained drivers who accepted the time discount. This raised the question of whether current courses are appropriate, and whether there are more appropriate approaches that could be adopted for improving novice driver safety. The current study involved a systematic review of the driver training literature to identify successful and unsuccessful approaches internationally. The review identified 82 relevant documents reporting a diverse range of programmes for novice drivers, which mainly focused on improving vehicle skills handling, reducing overconfidence, and/or developing hazard perception and situation awareness skills. Time discount policies have been used in some overseas jurisdictions, and even in the absence of explicit time discounts, driver training has often led to earlier licensure (eg, because the young driver feels better prepared to apply for their licence sooner). Research has consistently indicated driver training has a neutral or even negative effect on safety if it results in earlier licensure, with nearly all experts recommending against time discounts. Hazard perception training shows promise, but has predominantly been evaluated in small-scale, short-term experimental studies, so it is unclear if the benefits would transfer to real-world driving in the long term. Based on the review, we recommend that the current policy of providing a time discount for driver training should be discontinued. The poor quality of most driver training studies makes it difficult to identify best practice, as most evaluated programmes have substantial limitations, but we suggest some guidelines for future programmes.

1 Introduction

1.1 Background

Young drivers aged 16–24 are over-represented in serious and fatal crashes in New Zealand. Although drivers aged under 25 represent only 9% of current licence holders, in 2019 they were the primary at-fault driver in 26% of fatal crashes, 28% of serious injury crashes and 30% of minor injury crashes (Ministry of Transport 2021). The reasons underlying this over-representation are complex, with both inexperience and youth contributing to high crash rates (McCartt et al 2009).

New Zealand currently has a three-stage graduated driver licensing system (GDLS), in which drivers must first obtain a conditional learner licence, and then a restricted licence, before being able to progress to a full or unrestricted licence.

Individuals can apply for a **learner licence** if they are at least 16 years old. To obtain the learner licence, they must complete a computer-based multiple-choice theory test with at least 32 of 35 questions answered correctly. The learner licence allows individuals to drive under the supervision of another driver, who has held a full New Zealand driver licence for at least 2 years. Learner drivers must display 'L' plates and cannot drive after consuming any alcohol if they are under 20 years old.

Learners can apply for a **restricted licence** if they are at least 16.5 years old and have held their learner licence for at least 6 months. To obtain the restricted licence, they must pass a 45-minute practical driving test, which is completed in a vehicle arranged by the learner. A restricted licence permits the driver to drive on their own between 5 am and 10 pm, and with a supervisor between 10 pm and 5 am. Passenger restrictions apply when driving unsupervised, and drivers under 20 years old must not drive if they have consumed alcohol.

The minimum required duration for the restricted licence varies depending on the driver's age and whether they have completed an approved driver training course.

- Drivers under 25 years of age must hold their restricted licence for at least 18 months and must be at least 18 years of age before applying for a full licence. However, if they have completed an approved driver training course, the minimum duration is reduced to 12 months, meaning the minimum age for applying for a full licence is 17.5 years.
- Drivers aged 25 years and over must hold their restricted licence for at least 6 months but can reduce this requirement to 3 months if they have completed an approved driver training course.

Advanced driver training courses are designed to teach skills that can improve vehicle handling and hazard perception, which could theoretically help reduce the likelihood of being involved in a crash. This is why graduates of approved advanced driver training courses are permitted to apply for their full licence earlier than drivers who have not completed an approved course. Similar 'time discount' policies for driver training graduates have been adopted in overseas jurisdictions, although the specifics differ. For example, some jurisdictions apply the time discount to the learner stage instead of the restricted stage (Hirsch et al 2006), whereas others allow driver training students to obtain their licence at an earlier age (eg, at 15.5 years instead of 16 years; Mayhew et al 2014) but still require the same duration of licensure.

Time discount policies are intended to improve safety by encouraging individuals to complete advanced driver training courses, which are assumed to confer a safety benefit. However, a recent evaluation of the New Zealand GDLS indicated that crash rates were not significantly different when comparing trained and untrained drivers (Schiff 2019). The evaluation noted that most of the trained drivers who completed approved courses took advantage of the time discount.

When examining the subset of drivers who completed the advanced driver training course but did not accept the time discount, Schiff's (2019) evaluation found that crash rates were significantly *lower* compared with drivers who had not been recorded as having completed training. Specifically, the odds of crash involvement were 45% lower for drivers who completed advanced driver training but did not accept the time discount when compared with drivers without an advanced driver training completion certificate. This suggests that the potential safety benefits of driver training were counteracted by earlier licensure for drivers who accepted the time discount. This is consistent with decades of international research, which has suggested that accelerating the licensure process can negate the effects of training – with some programmes even being associated with increased crash risk as a result (eg, Lund et al 1986). It is also consistent with more recent research, which suggested that training is unlikely to improve safety among drivers who complete training with the aim of receiving an external incentive, such as a time discount or a reduced insurance premium (Hirsch et al 2006).

Although there is an extensive body of research evaluating driver training, the literature is diverse both in terms of its quality and its findings and implications (Beanland et al 2013). Some studies find that training does not significantly influence safety outcomes, as in Schiff's (2019) report, whereas others find that crash rates significantly *increase* post-training (ie, poorer safety outcomes), and some find that training improves safety. There are many reasons for these diverse outcomes, including variation in the content of training programmes, characteristics of the drivers who undertake training, and quality of the evaluations.

1.2 Aims and objectives of the research

The purpose of the research reported here was to investigate the approaches used currently in advanced driver training and review the effects that different training approaches have on drivers' skills, safety, behaviour and attitudes. A systematic review of the driver training literature was undertaken.

The objectives of the research were to:

- conduct a review of the national and international literature on advanced driver training
- compare international examples of effective advanced driver training across jurisdictions to identify key factors contributing to their success, and how and why those factors have been selected
- determine which elements of advanced driver training programmes are likely to contribute most to improvements in driver skill and behaviour.

The ultimate aim was to use the findings of the systematic review to make recommendations about which approaches are most likely to contribute to improvements in driver skill, behaviour and safety, and which approaches are likely to have lesser effects (or potentially negative effects – ie, resulting in poorer safety outcomes).

1.3 Details of this project

This project was conducted in New Zealand over the period November 2020 to March 2021.

2 Method

2.1 Overview of research approach

The research involved a systematic literature review of relevant driver training research, which followed procedures established by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al 2009), as detailed in this chapter.

Many systematic reviews also include meta-analyses, which involve statistical synthesis of quantitative results from multiple studies. However, a meta-analysis would not be feasible for the driver training literature given the wide variation of approaches taken in the literature. This includes diversity in the types of training programmes, the outcome measures studied (eg, crash rates, citations and violations, self-reported behaviour, hazard perception test scores, eye movements) and the evaluation timing (which ranges from literally within minutes of training completion, to longitudinal studies that evaluate drivers several years post-training). A further issue is that many studies provide incomplete or inconsistent reporting of statistics; for example, not reporting the exact number of individuals included in an analysis, or reporting only relative changes in crash rates without providing the actual number of crashes. These attributes of the driver training literature mean there would not be a sufficiently large number of sufficiently similar studies to warrant a meta-analysis.

2.2 Information sources

Three scholarly databases were searched:

1. PsycINFO
2. Scopus
3. Transportation Research Information Database (TRID)

The databases PsycINFO and Scopus were included to capture peer-reviewed academic literature (ie, journal articles and conference proceedings) as well as book chapters.

TRID was included to capture more diverse sources that focus specifically on transport, which may not be indexed by other databases, as well as unpublished or 'grey' literature such as government and consultant technical reports (which is a common method for publishing driver training evaluations).

2.3 Search strategy

The database searches were conducted on 16 November 2020 with no date restrictions applied. Although it was assumed that more recent studies would be more relevant, due to changes in licensing approaches (eg, introduction of a GDLS in many jurisdictions) and other improvements to road safety, we did not automatically exclude older studies. We made this decision because there are some older studies that involved large-scale evaluations of training programmes, which can provide useful knowledge regarding what works and does not work in driver training.

Each database was searched using 'driver training' and 'driver education' as keywords. Although the project explicitly focuses on driver training, it is important to include driver education in initial searches as many people use the terms 'training' and 'education' interchangeably in this context.

The searches were not specifically restricted to focus on young or novice drivers, for two reasons. One was to avoid accidentally missing relevant studies by using search parameters that were too narrow. The second reason was to allow inclusion of appropriate training programmes that have only been evaluated using more experienced drivers, but which could be appropriately marketed to young novice drivers.

Five separate database searches were conducted (see Table 2.1).

- For Scopus and TRID, two searches were conducted: one using the keyword term ‘driver training’ and one using the keyword term ‘driver education’. This approach of searching the databases twice yielded some duplicate results but minimised the potential for missing relevant items.
- For PsycINFO, which includes a database-specific subject heading ‘Driver Education’, a single search was undertaken to capture all items that were either indexed under the ‘Driver Education’ subject heading, or otherwise included the term ‘driver training’ as a keyword.

This search process yielded 4,472 search results. The results of each search were exported into a custom database created using EndNote, a software program for managing bibliographies, citations and reference lists. The information exported to EndNote included all available bibliographic details about the reference, as well as its abstract (where available).

Table 2.1 Database searches, including search terms, limits applied and results returned

Database	Search term	Limits applied	Results
PsycINFO	‘Driver Education’ (subject heading) OR ‘driver training’ (search as keyword)	Language: English	519
Scopus	‘driver training’ (in article title, abstract, keywords)	Language: English Subject areas: Engineering; Social Science; Medicine; Psychology	1,926
Scopus	‘driver education’ (in article title, abstract, keywords)	Language: English Subject areas: Engineering; Social Science; Medicine; Psychology	481
TRID	‘driver training’	Language: English Subject areas: Education and Training; Policy Result type: All publications	761
TRID	‘driver education’		785
		Total search results	**4,472**

2.4 Results screening

After the search results were exported into EndNote, items were screened to remove duplicates and out-of-scope documents. This process is summarised in Figure 2.1.

2.4.1 Duplicate items

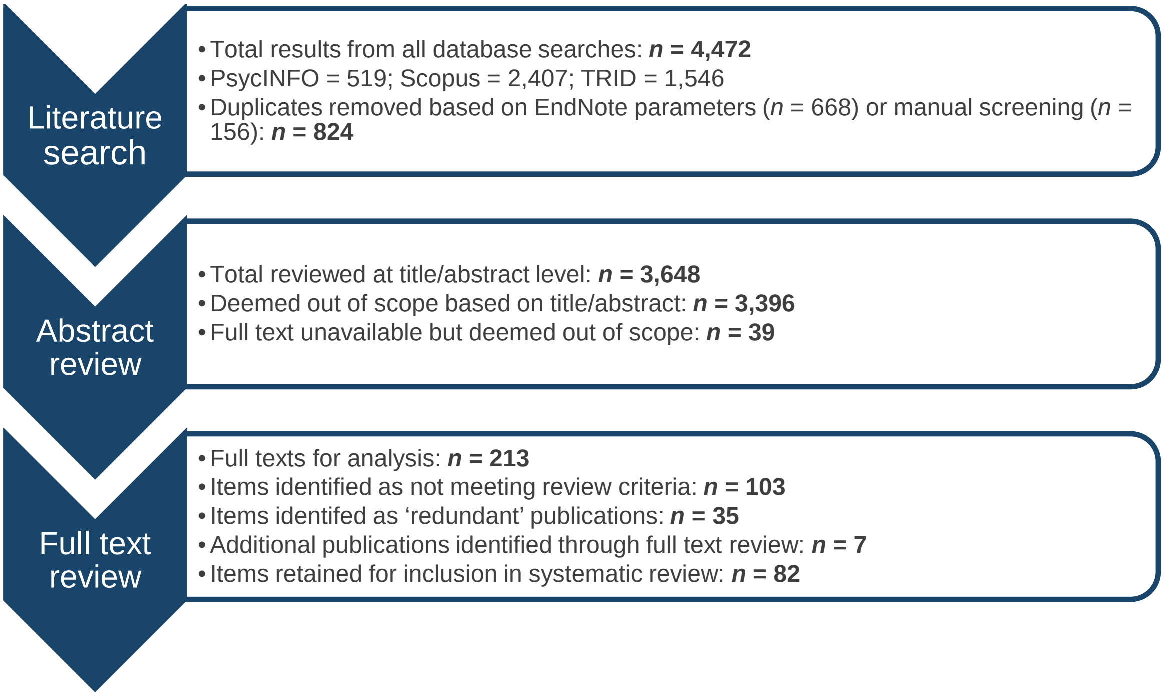
Because scholarly databases share some overlap in the items they index, some of the search results were duplicates.

EndNote has a ‘Find Duplicates’ function for automatically identifying duplicates, which compares references based on the title, author and year of publication. Using this function, 668 items were removed as duplicates.

Some duplicate items are not identified using the automatic function, because of slight differences in the information entered in different databases. For example, different databases may enter authors’ names slightly differently (full name, single initial, all initials). In other cases, records may contain inaccuracies that result in differences between databases. To identify additional duplicates, items were manually screened by a researcher, comparing titles and authors. This manual process led to a further 156 duplicate items being removed.

The processes described above removed items that were exact duplicates (ie, the same paper listed multiple times). Closer review of the items revealed another type of 'duplicate', whereby the same dataset was reported in multiple publications. This type of duplicate publication can occur because researchers disseminate their findings in multiple formats. Within transport research, often researchers will write a technical report (if the work was undertaken for government), then present at a conference, which may require writing a conference proceedings paper (5–10 pages), and then may also write up the study for publication in a journal article and/or book chapter. If included within a systematic review, these items can bias the results by creating a false sense of consensus – because there will be multiple studies that reach (presumably) the same conclusions. These types of 'duplicate publications' were identified at a later stage, while reviewing abstracts and full text papers.

Figure 2.1 Systematic literature review process, including items excluded at each stage



2.4.2 Title and abstract screening

After removing duplicate records, the remaining 3,648 records were screened for relevance. The first stage of screening involved reviewing titles and abstracts to remove any items that were obviously out of scope. Most items were rejected as out of scope based on the title and abstract review (*n* = 3,396).

Key reasons for identifying a paper as out of scope included:

- it was not transport focused (ie, completely unrelated topics)
- it focused on transport modes other than passenger cars (including railways; motorcycles and mopeds; bicycles; pedestrians; and buses, coaches, trams, streetcars, and other forms of mass transportation)
- it focused on work-related driving (including taxi drivers, truck drivers, other commercial drivers, and fleet drivers)

- it focused on subpopulations that were deemed out of scope, including older adults, children too young to drive, parents of teen drivers, and transport users with a disability or medical condition
- it focused on an aspect of road safety other than driver training (eg, factors associated with crash risk, aspects of the GDLS unrelated to driver training).

Some items that focused on driver training or driver education were excluded because they were assessed as being out of scope or unsuitable for inclusion in the systematic review. These fell into two broad categories:

1. Specialised driver training. Examples included training drivers to use specific advanced driver assistive systems or automated vehicles; training for drivers with a specific disability or medical condition, such as those recovering from stroke; and court-mandated training for offenders. These programmes were excluded because the aims and design of training would not be relevant to the broader population of young novice drivers.
2. Items that did not include a relevant empirical evaluation. These included previous literature reviews of the efficacy of driver training (eg, Beanland et al 2013) and opinion pieces describing what driver training should involve.

Items were retained for further review if there was any ambiguity about the paper content (ie, if the paper appeared to be about driver training but it was unclear whether the paper reported an evaluation).

In addition to the items rejected after title and abstract review, a further 39 items were removed from consideration because they did not have an accessible full text. Most of these were older publications (20–75 years old), and after scrutinising the abstracts and other available information (eg, other publications in the same outlet, by the same authors, or subsequent papers on a similar topic) it was determined that these items were not likely to be relevant to the review.

2.4.3 Full text review

Following the title and abstract screening stage, 213 items remained for full text review. Items were retained if they met the following criteria:

- They involve a training programme that is designed to increase driver skills and/or safety.
- The authors report some data about the programme's outcomes that are relevant to on-road behaviour and safety (ie, post-training skill, behaviour, attitudes, crashes, or violations).
  - Items were excluded if they focused on process evaluation (ie, participants' perceptions of the course and whether they enjoyed it or considered it worthwhile).
- The training is intended for car drivers.
  - Evaluations of training programmes for other road users were excluded, including programmes that focused exclusively on 'pre-drivers' (ie, young people who are not yet old enough to apply for a licence) or parents of young drivers.
- The training is intended or appropriate for novice drivers progressing through the learner or restricted phase of a GDLS.
  - Evaluations were excluded if they focused on skills that are appropriate for more advanced drivers (eg, off-road driver training).
  - Some evaluations were retained if they focused on experienced drivers but involved training that could be appropriate for learner or restricted drivers.

A total of 103 items were removed from consideration because they did not meet the review criteria.

At this stage, items were also excluded if they were deemed 'redundant publications' – that is, publications that reported a dataset that was analysed in a separate publication, or a pilot of a larger study that was subsequently published. On this basis, 35 publications were removed from consideration. This included some longitudinal studies, as the authors reported on the dataset at multiple time points. It also included some conference papers, as in some cases researchers submitted the same study to multiple conferences as well as publishing a journal article and/or technical report. Where redundant publications were identified, the paper with the most complete dataset was retained for the review, and related publications were consulted to confirm relevant details if necessary. For longitudinal studies, the paper reporting the longest follow-up period was considered as the 'most complete dataset', even if it did not contain the largest number of participants, because a key aim of the review was to identify effects that would have a long-term influence on safety. For non-longitudinal studies, the paper with the largest number of participants was included. If there were two or more papers that reported data for the same number of participants, the paper with the most relevant outcome variables was included.

These screening and selection processes resulted in 75 items being retained. An additional 7 relevant publications were identified through reviewing full texts (ie, by reviewing references lists), resulting in a final sample of 82 publications being included in the review.

2.4.4 Data extraction

The following data were extracted from publications that were included in the systematic review.

- The nature of the driver training programme/intervention delivered, focusing on:
  - programme duration
  - skills targeted (eg, vehicle handling, hazard perception, knowledge, insight)
  - training format and mode of delivery (eg, on-road, test track, simulator, computer-based, classroom, online, self-paced)
  - the characteristics of the sample used to evaluate the training programme (eg, number of drivers, age, driving experience and licensing stage (equivalent to New Zealand GDLS stage), and any relevant recorded sociodemographic characteristics).
- Evaluation characteristics, focusing on:
  - whether the study used a within-subjects (pre- vs post-training comparison) or a between-subjects (trained vs untrained) design, or some other comparison
  - whether participants were randomly allocated to conditions (for between-subjects comparisons)
  - timing of the evaluation (ie, immediately post-training, short-term follow-up, long-term follow-up)
  - outcome variables recorded (eg, crash rates, violations, real-world driving performance, simulated driving performance, hazard perception or anticipation, self-reported behaviour)
  - effect of training on outcome variables.

Issues relating to the quality of training programmes and evaluations were noted where possible; for example, if there were obvious confounds or limitations, or if information was missing, unclear or obviously erroneous (eg, typographical errors in publications). These issues were noted to assist in reconciling conflicting findings. Where two similar studies yield conflicting results, it makes sense to put greater weight on the findings from a higher-quality, more robust evaluation. The quality of each report was not formally scored, largely because the issues varied greatly between studies and assigning an overall score would have obscured this variation. For example, some studies included large samples but involved non-random assignment to conditions, which means training is confounded with other factors (some of which are unknown or unmeasured). Other studies used randomised assignment with well-matched samples, but the sample sizes were relatively small and/or the evaluation was only conducted immediately after the training, so the long-term effects of training are unknown. It is important to consider the specific limitations of each study individually and the implications of any limitations, in terms of whether these results are likely to generalise to other contexts.

3 Results

Because of the large number of papers identified in the systematic review, papers were grouped into broad categories based on the type of training evaluated.

Given the motivation for the research, more detail is provided about programmes that are linked to the licensing process and/or incorporate on-road training, as these are the most comparable to the programme currently eligible for a time discount in New Zealand.

3.1 Pre-licence training within the licensing process

Several jurisdictions incorporate pre-licence training within the licensing process (or have done so in recent decades), with the training programmes either created or explicitly endorsed by the government agency responsible for driver licensing.

Programmes were classified in this category if they were aimed at learner drivers and were intended to help the learner achieve their restricted or full licence (or international equivalent). Broadly construed, these programmes can be sorted into three subtypes:

1. compulsory driver training for licence applicants
2. driver training as an alternative pathway to licensure (ie, the jurisdiction offers different learner pathways, one of which involves undertaking a specific training programme)
3. driver training as a ‘shortcut’ to licensure (ie, graduates receive a time discount, or can obtain their licence at an earlier age).

None of the studies summarised in this section involve random assignment to the training condition, as they involve real-world training programmes that have been implemented on a large scale, with self-selection and/or naturally occurring groups. Therefore, where there are comparisons between trained and untrained drivers, there may be systematic differences between groups that could either exacerbate or obscure the effects of training.

3.1.1 Finland: Compulsory skid training

Finland adopted a new model of driver licensing and training in 1990, which included compulsory skid training. Although previous studies in the 1980s and 1990s had suggested skid training increases crash risk, this module was deliberately designed to emphasise risk awareness rather than vehicle manoeuvring; for example, avoiding slippery roads altogether.

Katila et al (2004) evaluated the introduction of compulsory skid training in Finland using data from two surveys. One survey provided self-reported crash data for approximately 30,000 drivers licensed in 1989 (‘old curriculum’ or untrained group) and 1990 (‘new curriculum’ group that received compulsory skid training). The second survey asked a subsample of nearly 1,300 drivers to evaluate their driving skills. Crash data for newly licensed drivers were also compared with broader population crash trends (ie, police-reported slippery road crashes) over the same period.

The evaluation revealed that, when accounting for broader crash trends (ie, population changes in slippery crash rates) that the rates of slippery road crashes neither increased nor decreased significantly (Katila et al 2004). The new skid training curriculum increased drivers’ confidence in their slippery driving skills, even though it was designed not to. The fact that this was not associated with increased crash rates highlights that the relationship between confidence and safety is complex.

3.1.2 Denmark: Compulsory driver training

Denmark introduced a new compulsory driver training curriculum in 1986 with linked theoretical and practical components, such that learners had the opportunity to put skills into practice shortly after learning about it in class (Cartensen 1994). Both the new and old programmes of driver training restricted supervised practice to sessions with a qualified driving instructor (ie, no supervised practice with family or friends). The key difference was that the new curriculum was more structured and went beyond teaching traffic laws and basic car control, to emphasise defensive driving and hazard perception.

Cartensen (2002) evaluated the new curriculum, first by comparing young driver crash rates before and after the change. Analysis of police-reported injury crashes showed that crashes for all drivers decreased over time, but the relative decrease for drivers aged 18–19 (ie, those affected by the law) was approximately 20% greater than the reduction in crash rates among older, more experienced drivers.

In the second stage of the evaluation, Cartensen (2002) surveyed approximately 4,000 drivers who were licensed either before or after the new curriculum was introduced, to capture respondents' self-reported driving exposure and crashes over their first 5.5 years of licensure. In their first 18 months of licensure, drivers trained using the new curriculum had significantly lower crash rates when adjusted for exposure. Multiple-vehicle crashes reduced by 17%, and manoeuvring crashes reduced by 31%, but there was no significant difference in single-vehicle crash rates. Cartensen suggested that the new programme of driver training may be more effective at targeting factors underlying multiple-vehicle crashes, rather than single-vehicle crashes, because it addressed skills such as vehicle handling and hazard perception.

There were no significant differences in crash rates between the two groups when examining the period from 18 months to 5.5 years post-licensure, but crash rates declined for both groups.

Importantly, some students received the new curriculum before it became compulsory in 1986, and these drivers also had lower post-licensing crash risk, suggesting the reduction was associated with the structure and content of the training, rather than being an artefact of time passing.

3.1.3 Quebec, Canada: Compulsory driver training (1983)

On 1 January 1983, the Canadian province of Quebec made it compulsory for all new drivers to complete an approved driver training course through a registered driving school. The courses included 30 hours of classroom instruction and 8–10 hours of in-vehicle driving instruction. Prior to 1983, driver training was compulsory for 16- and 17-year-old drivers only.

To examine the effect of making driver training mandatory for all ages, Potvin et al (1988) analysed changes in crash rates for newly licensed drivers in 1983–84 compared with 1980–82, using drivers aged 18+ as the training 'treatment' group and drivers aged 16–17 as the comparison group. The authors found different patterns for male versus female drivers, so the analyses were split by gender.

Among males, there were no significant changes in crash rates for either the treatment or control group, and no significant change in licensing rates. Before 1983, most males obtained their driver's licence before the age of 18 and therefore most males were already required to complete a training course, so mandating driver training did not have a significant effect.

Among females, the treatment group showed no significant difference in injury crash rates, and no change in non-injury crashes in the first 6 months of licensure. In the second 6 months of licensure, non-injury crashes significantly increased for females in the treatment group. However, the comparison group also showed significant increases in all crashes – that is, mandating driver training for drivers of all ages appears to have significantly *increased* crashes among 16- and 17-year-old females.

This increase in crashes likely occurred because the new law was associated with a statistically significant 19.5% increase in the number of 16- and 17-year-old females obtaining a license. Prior to 1983, females' modal age of licensure was 18 because the convenience of earlier licensure was offset by the inconvenience of having to undertake training. When the law changed to also require 18-year-olds to undertake training, there was no remaining advantage to waiting to get a licence, thus effectively encouraging earlier licensure.

Overall, these findings do not suggest a clear benefit of the training, as there was no reduction in crashes, and changing the law seemed to encourage earlier licensure for females corresponding with increased crash rates in this group.

There are two major caveats to these findings.

1. It is unclear how many adults completed driver training prior to 1983, but Potvin et al (1988) estimated 60–70%. Therefore, the comparison is not simply between 'trained' versus 'untrained' drivers; rather, it is between a group that always had training (ie, the comparison group of 16- and 17-year-olds) and a group for whom the rate of training increased from 'partial' (possibly 60–70%) to 'complete' (presumed 100%). This makes the results more confusing conceptually.
2. The 95% confidence intervals for the effects are extremely wide, indicating the estimates are unreliable. This study presumably involved all newly licensed drivers in Quebec (although the number of individuals included in the analysis is never stated), so this lack of statistical reliability cannot be attributed to sample size. The unreliable findings are more likely to be because crashes are relatively rare events, so it is hard to identify factors that reliably change crash rates. (More common events have more data available, so it is easier to identify trends and predictors.)

3.1.4 Quebec, Canada: GDLS time discount (2000–03 evaluation)

A subsequent change to Quebec's GDLS made driver training optional, with drivers who completed approved training eligible to reduce the minimum time on their learner's permit from 12 months to 8 months. Approved training courses were supposed to involve at least 12 hours of practical lessons, with an optional theory module.

Hirsch et al (2006) evaluated this policy in a study of 1,804 respondents, of whom 85% had completed training. There was no association between training and either violation or crash rates in the first 450 days of licensure. Taking the optional theory module was also not associated with subsequent crash rates (but was associated with increased likelihood of passing the theory test).

Learners' self-reported motivations for undertaking driver training were grouped in three categories: to improve learning; for opportunity (ie, time or insurance discounts); or both learning and opportunity. Motivation predicted both violation and crash rates post-training. Drivers who did training for opportunistic reasons had 40.7 violations and 14.8 crashes per 100 drivers, whereas those who did it for learning had significantly lower rates of 14.0 violations and 5.7 crashes (with the 'mixed' group in between).

Training was also associated with drivers completing fewer hours of driving practice prior to licensure. Students who completed the training were significantly more likely to report < 25 hours of supervised practice (ie, with a parent or guardian), which combined with the training would total < 37 hours on-road, whereas drivers who did not complete training were more likely to complete > 50 hours.

There were also apparent demographic differences between the trained and untrained drivers, with those who completed training reporting greater financial support from family (suggesting higher socio-economic status).

3.1.5 Ontario, Canada: GDLS time discount

Additional studies evaluating the impact of 'time discounts' at the learner stage of the GDLS have come from Ontario, Canada, in the late 1990s. Ontario's GDLS involved two licence stages before full licensure – G1 and G2 – with both stages requiring a default period of at least 12 months. The G1 stage is closest to the New Zealand learner licence and requires supervised driving, with some other conditions. The G2 stage is most similar to New Zealand's restricted licence (albeit with fewer restrictions at the time of the evaluations reported). Learners could reduce the G1 stage to 8 months by taking an approved training course that included a minimum of 25 classroom hours and 10 hours of driving practice.

Initial analyses after the introduction of this policy indicated that driver training was associated with significantly higher crash risk (44% higher) during the G2 stage (Boase and Tasca 1998). However, this analysis did not factor in differences in exposure or driving experience between groups.

To account for this, Zhao et al (2006) analysed survey data from 1,528 high school students aged 16–18, of whom two-thirds had completed driver training. The duration of follow-up varied between respondents, depending on when they received their licence, but could be up to 20 months post-training. During the G1 (learner) stage, odds of a self-reported collision were significantly lower for drivers who had completed training compared with those who had not. Two points should be noted here: first, crashes are relatively rare during the supervised learner phase (only 3% reported having a collision); second, completing training reduced the time spent at the learner phase, which effectively meant trained G1 drivers had less opportunity to crash.

During the G2 phase, there was no significant difference in self-reported crashes when comparing trained and untrained drivers in a multivariate regression that accounted for demographic characteristics, driving experience and exposure.

3.1.6 Netherlands: Intensive versus distributed training

Research from the Netherlands has compared the effects of intensive versus distributed training for young drivers. In this system, learners can only practice driving with a qualified instructor. Most drivers take a traditional approach of approximately 1 hour per week, which requires several months to prepare for licensure (de Craen and Vlakveld 2013).

De Craen and Vlakveld (2013) evaluated the long-term impact of an intensive course in which learners practice for several hours a day over a much shorter period (eg, 2 weeks). They surveyed 436 newly licensed drivers, of whom 39 (9%) had completed the intensive course, from 2 weeks to 2 years post-licensure. Obviously, participants self-selected into training condition, but the two groups were not significantly different on any demographic variables (age, gender, education, urban/rural) or driving habits (exposure, alcohol, speeding, seatbelt use) that were measured. The intensive group were significantly more likely than the distributed group to report 'incidents' (defined as any event resulting in property damage or injury) in the first 2 years of licensure (43% vs 26%). The intensive group also reported completing fewer lessons than the distributed group but did not differ on hours of driving practice (ie, they took longer lessons) or pass rates for their licence test.

De Craen and Vlakveld's (2013) evaluation suggests that distributed driving practice is superior for long-term skill retention, which is consistent with broader psychological literature on distributed versus massed (intensive) practice. However, the sample size is relatively small, and the evaluation lacks details about both the types of crashes and the types of situations encountered in training.

3.1.7 France: Supervised driving apprenticeship

France has two pathways for driver licensing. The ‘traditional’ route was to take lessons with a driving instructor to prepare for their licence test. The alternative route – l’apprentissage anticipé de la conduite (AAC) – allows learner drivers aged at least 16 years to complete a 20-hour practical driving course, followed by a 1–3-year period of supervised practice. During the supervised driving ‘apprenticeship’, students are required to drive 3,000 km supervised by any qualified driver (typically a parent).

In a case-control study of 1,145 young drivers, Page et al (2004) found that the type of initial training (AAC vs traditional) did not significantly predict crash likelihood in the first 2 years of licensure.

AAC training was associated with early licensure, as 90% of AAC students got their licence at age 18 (rather than 19 or 20, which is common in France) whereas only 50% of traditional students did. However, AAC-trained drivers had significantly lower exposure (less driving) in their first 3 years of licensure, which may have offset the risk of early licensure.

Analysis of AAC drivers’ training logs revealed that they drove on average 5,000 km, much more than the minimum requirement, but crucially the practice was not diverse and primarily involved ‘easy’ driving. Page et al (2004) suggested supervised practice conditions need to be more varied and target more specific, potentially difficult scenarios (ie, the types of scenarios that may be problematic or challenging when driving unsupervised). They also hypothesised that spending so many hours driving with a parent or other regular supervisor could make learners over-reliant on the support, making it harder for them to transition to solo driving later.

It should be noted that the case-control participants were all young drivers who held a comprehensive insurance policy, meaning the sample is likely biased and not representative of the broader population, although it is unclear whether this bias would influence the impact of supervised driving practice.

Freydier et al (2016) also evaluated the French apprenticeship system using a small-scale simulator study of 42 drivers: 15 AAC novice drivers, 13 traditionally trained novice drivers and 14 experienced drivers (aged 25–35). AAC drivers were significantly better than traditionally trained drivers at speed regulation and maintaining lateral position. They were also better at performing a concurrent secondary task, suggesting they were more skilled at the primary task of driving. Experienced drivers were significantly better than AAC drivers at speed regulation, but other differences were not significant. These findings suggest that the AAC early training model can significantly improve vehicle handling skills. An obvious limitation is that this study used a simulator, rather than real driving, and involved a small sample that may not be generalisable to the broader driving population.

3.1.8 Nebraska, USA: Driver training vs supervised driving

In Nebraska, USA, teens can apply for a learner’s permit at age 15 and a provisional operator’s permit (POP) at 16. The POP is like New Zealand’s restricted licence: it includes night-time driving and peer passenger restrictions and must be held for 2 years before drivers can obtain a full licence.

To obtain a POP, teens could either undertake 50 hours of supervised driving practice (including 10 hours at night) or they could complete a certified driver education course, which included 20 hours of classroom instruction and 5 hours practical instruction.

Shell et al (2015) compared these two avenues for obtaining a POP in a population study involving 151,880 teens. The trained group were significantly lower in both overall crash rates (Year 1: 11.1% vs 12.9%; Year 2: 8.4% vs 9.5%) and casualty crash rates (Year 1: 2.3% vs 2.6%; Year 2: 1.5% vs 1.8%). Training was also associated with significantly lower rates of violations, with the magnitude of the difference being even larger (Year 1: 10.4% vs 18.3%; Year 2: 13.4% vs 20.9%).

There were systematic differences between the trained and untrained drivers, with the untrained (supervised practice) drivers more likely to be rural, male, non-white, lower socio-economic status and older teens. Training remained a significant predictor of both crashes and violations in a logistic regression that factored in these sociodemographic variables that influence driver risk. However, other factors may differentiate the groups, such as motivation to learn and/or drive safely.

This is one of the few large-scale training evaluations that found a significant safety benefit. It is notable that the training intervention does not provide a time discount, nor does it reduce the age of licensure. Training reduced the minimum required hours of supervised driving practice, from 50 hours for the supervised practice group to 5 hours for the driver training cohort. However, the training group were allowed to do additional supervised practice and were not required to log their hours, so it is unclear how many hours of supervised practice they accrued outside their driver training course.

3.1.9 Oregon, USA: Driver training vs supervised driving

In Oregon, USA, GDLS restrictions apply for drivers aged under 18. Teens can obtain a learner's permit at 15 and apply for their provisional (restricted) licence at 16. They must hold their learner's permit for at least 6 months and must complete 100 hours of supervised driving practice, but if they complete an approved driver training course, they are only required to complete 50 hours of supervised driving.

Approved training courses are offered through several providers, including high schools and driving schools. They must include a minimum of 30 hours of classroom instruction, 6 hours of on-road instruction, and 6 hours of in-car observation of another student. The training must be spread over multiple weeks to ensure distributed practice (eg, classroom time must be a maximum of 3 hours per day and 6 hours per week). Simulator driving can substitute for up to half of the on-road driving practice, at a 4:1 ratio (ie, 4 simulator hours substitute for 1 on-road hour).

Mayhew et al (2014) evaluated the training option, which was completed by just under one-third of Oregon teens. Their main evaluation involved a survey of over 4,200 recently licensed teens, including 1,000 who had completed the training programme. Violation rates were very low and did not differ significantly between the trained and untrained groups.

Officially recorded crash rates were significantly higher among teens who completed the training programme, compared with untrained teens. This effect appeared mostly due to increased crash rates among 16-year-old and male drivers who had completed the training. Driver training was associated with several sociodemographic variables, with trained drivers being significantly younger and more likely to be male, white and urban, with higher parental education and greater support for and knowledge of the GDLS. When sociodemographic variables were included in a multivariate analysis, training status no longer significantly predicted crash rates.

Following the survey, Mayhew et al (2014) examined historical records for over 94,000 teen drivers. The benefit of studying records is that it avoids sampling and response bias, which inevitably occurs in studies that require active participation. For example, drivers who respond to a survey about road safety may be inherently more safety conscious. However, fewer details about the drivers are available.

The historical records showed that 22% of teen drivers had completed training, and individuals in the training cohort were significantly younger than the untrained cohort (consistent with the survey results). When accounting for length of licensure, trained drivers had significantly lower crash rates in the first 18 months of licensure (but not beyond that time). In a multivariate analysis including sociodemographic variables, training remained a significant predictor associated with a 4.3% reduction in crash rates. Trained drivers also had significantly lower conviction rates throughout the first 54 months after obtaining a provisional licence, with untrained drivers having approximately twice the rate of convictions compared to trained drivers.

3.1.10 Manitoba, Canada: Earlier licensing

In Manitoba, Canada, driver education is optional and allows teenagers to obtain their learner licence 6 months earlier – at 15.5 years instead of 16 years old (Mayhew et al 2014). Notably, this means that driver education enables *earlier* licensing, but there is no ‘time discount’ because they must still hold their learner licence for at least 9 months.

Manitoba Public Insurance provides a high school based programme which is taken by approximately two-thirds of all students (Mayhew et al 2014). The programme includes 30 hours of classroom instruction, 8 hours of practical driving instruction and 8 hours of in-car observation of another student. Students must also complete an additional 24 hours of supervised driving practice within 9 months of completing the programme to obtain a graduation certificate. Thus, recognised graduates should accrue at least 32 hours driving on-road and 38 hours of other education and training.

Mayhew et al (2014) reported a series of studies evaluating Manitoba’s high school driver education, comparing trained and untrained drivers in a survey study, in a simulated driving test, and by examining their performance in their on-road intermediate licence test.

The survey had 1,050 respondents, of whom 59% had completed high school driver education. Self-reported risky driving was rare and did not differ significantly between trained and untrained drivers. The trained drivers gave significantly higher self-ratings of their driving skill, although the absolute size of this difference was small. Trained drivers also reported making fewer driving trips and having spent less time driving recently, which may be because they were more likely to live in major cities.

The simulator study compared driving errors among 170 teen drivers of varying experience and training status: 66 were pre-drivers; 63 were learner drivers, of whom 50 had completed driver education; and 41 were intermediate (restricted) drivers, of whom 37 had completed driver education. The number of errors made during simulated driving decreased as experience increased (ie, from pre-driver, to learners, to intermediate). Among learner drivers, those who had completed driver education made significantly fewer errors, particularly for errors involving speed, stopping violations, and at-fault collisions. Learner drivers who had completed driver education also self-reported significantly fewer risky driving behaviours.

There were no significant effects of driver education among the intermediate drivers; however, the ‘no education’ group was far too small (n = 4) to permit any meaningful conclusions from this simulator study.

A significant limitation of Mayhew et al’s (2014) reported simulator study is that the dependent measure was driving errors, which had to be coded by observers. Coding is subjective, so the reliability and consistency of coding must be evaluated. In this case, the kappa statistics representing inter-rater agreement were < 0.4, indicating ‘poor’ agreement between different coders. This suggests that different coders were adopting different criteria for designating driving behaviours as errors. The authors downplay the impact of this issue.

The final aspect of Mayhew et al’s (2014) evaluation examined the performance of 2,935 teenagers aged 16–19 applying for their intermediate licence; 86% of whom had completed driver education. The driver education group were significantly younger, with an average age of 17 years 1 month, compared with 17 years 11 months for the comparison group.[1] The pass rate was 52% and did not differ significantly as a function of driver education. However, scores were significantly higher for applicants who had completed driver education, both among those who passed and those who failed. This suggests that driver education

[1] The minimum age for applying for an intermediate licence in Manitoba is 16 years 3 months for those who have completed approved driver education and 16 years 9 months for those who have not.

graduates may be slightly better prepared for the licence test, although it is unclear whether this would translate into safer outcomes during licensure.

3.2 Post-licence training programmes within the licensing process

Only one paper (Begg and Brookland 2015) identified in the current systematic review involved evaluation of a post-licence training course that is explicitly linked to the licensing process (although, as noted in subsequent sections, numerous *optional* post-licence training courses have been evaluated). This paper focused on New Zealand advanced driver training courses.

Begg and Brookland (2015) evaluated the impact of participating in defensive driving courses as part of the New Zealand Drivers Study, a longitudinal cohort study of young drivers. They interviewed 1,763 drivers, of whom 49% had participated in a defensive driving course. Only 1% had participated in an alternative course, so their analyses focused on defensive driving.

Most drivers who completed the course accepted the time discount (86%) and admitted this was their reason for taking the course (85%). Those who accepted the time discount had a significantly higher risk of receiving a traffic offence notice on their full licence (relative risk = 1.4, or 40% higher). The relative risk increased when driving-related variables were added to the model but decreased when demographic variables and time on full licence were included; however, the risk remained significantly higher (relative risk = 1.1, or 10% higher).

The driver training participants were younger at full licence (unsurprising, as most accepted the time discount), tended to reside in low-deprivation (high decile) areas, and were less likely to be Māori. There were no significant differences in gender, personality traits, or self-reported alcohol and drug use.

3.3 Courses incorporating on-road training

Several studies evaluated driver training courses that incorporate both practical lessons and classroom or discussion-based sessions. These are similar to the programmes described in previous sections that are linked to various licensing systems; however, being optional, most programmes incorporating on-road training are designed for the post-licence phase.

Post-licence courses are usually shorter than pre-licence courses; whereas many pre-licence courses comprise 20–50 hours of training, post-licence courses often last only 1 day (or the equivalent of 1 full day, spread over multiple days).

In Europe, post-licence training courses are often referred to as ‘second phase’ courses because the learning process is conceptualised as having two phases: pre-licence (often mandatory training) and post-licence (typically optional).

3.3.1 Georgia, USA: DeKalb County pre-licence training

One of the most widely discussed studies of driver training was conducted in DeKalb County in Georgia, USA, in the late 1970s and early 1980s. The DeKalb study involved a randomised controlled trial with over 16,000 high school students randomly assigned to a control group or one of two training programmes: Safe Performance Curriculum (SPC) or Pre-Driver Licensing (PDL).

The SPC programme totalled 70 hours of instruction, including 32 classroom hours with the remaining time including simulation, driving range, evasive manoeuvring, and on-road driving. The PDL programme included similar components but only involved 21 hours of instruction. The amount of on-road time was low for both programmes (1 hour for PDL; 3 hours 20 minutes for SPC). Details of the curriculum are unknown, but it is presumably outdated, given changes in vehicle technology, road infrastructure and the introduction of the GDLS. However, it is worth mentioning this study because it highlights the challenges in evaluating driver training.

Initial analyses of the DeKalb study data were provided in a technical report by Stock et al (1983). This report concluded that driver training significantly improved skills, with SPC students outperforming PDL students, who outperformed controls. Trained drivers had fewer crashes than controls during their first 6 months of licensure, with no difference in the subsequent 18 months. Violations were significantly lower among trained drivers than controls for the first 12 months of licensure, but not the second 12 months.

The analysis by Stock et al (1983) used raw, unadjusted crash and violation rates. Lund et al (1986) argued this was inappropriate because of systematic differences between groups, particularly in terms of exposure and licensing rates. They therefore reanalysed the data to compare crash risk adjusted for exposure. This reanalysis found that SPC students had higher rates of both crashes and violations than controls, whereas PDL students were not significantly different to controls. Both driver training programmes were associated with earlier licensure, and higher rates of licensure, suggesting that any potential benefits of the programme (ie, improved skills and knowledge, as measured by Stock et al 1983) were undermined by accelerated licensure.

These conflicting conclusions highlight the importance of carefully planning evaluation studies (and training programmes, as the SPC programme was never properly piloted), measuring confounding variables, and ensuring statistical analyses are appropriate. It is a useful example in that two analyses of the same dataset reached opposing conclusions because they adopted different analytical approaches. Specifically, Stock et al (1983) simply compared outcomes between training groups, whereas Lund et al (1986) used statistical analyses that factored in differences in licensure and exposure between groups. This should be kept in mind when considering the conflicting findings of other studies, because different studies take slightly different approaches to data analysis and adjusting for potential confounding variables. Even where two studies consider the same confounding variable (eg, exposure or socio-economic status), it may be measured differently, which could also influence the results.

3.3.2 Maryland, USA: Skid avoidance training

Farmer and Wells (2015) evaluated an optional skid training that is available to young drivers in Maryland, USA. As part of the Maryland GDLS, learner drivers must provide evidence that they have completed an approved driver education course (similar to the courses described in section 3.1) before they can apply for their provisional licence. Following this, drivers can take an optional 90-minute skid avoidance module. Farmer and Wells recruited 234 teen drivers to take the skid training and compared them with 5,046 similar stage teen drivers who had not completed skid training.

Of the prospective participants that were invited to undertake the training, only 16% accepted, raising concerns about the representativeness of this sample. To account for this, Farmer and Wells (2015) ran multiple analyses comparing the training group with different subsets of controls (eg, matched on key demographic characteristics). Unfortunately, these analyses resulted in inconsistent results, with some comparisons showing no significant effect of training, but others showing variously that training significantly increased violation rates, significantly increased crash rates, and significantly decreased crash rates. Given this, the authors concluded that skid training did not have a clear effect on safety. This study highlights the problem of trying to plan appropriate analyses when the training group is systematically different to the control group and/or not representative of the broader young novice driver population.

3.3.3 Sweden: Insight vs skills pre-licence skid training

Skid training has been a particular focus in Nordic countries, including Sweden, because novice drivers need to learn to negotiate icy and slippery roads. Gregersen (1996) compared two strategies for training drivers to handle skids safely. The 'skill' group had traditional skid training, which focused on vehicle handling skills, and drivers were explicitly instructed in how to operate the pedals for optimal performance. In contrast, the 'insight' group were given only a brief overview of the theory of driving and braking on icy roads but were left to make their own decisions, and the training was designed to emphasise the unpredictability and uncontrollability of skid situations.

Fifty-three learner drivers aged 18–24 were randomly assigned to one of the two training conditions and were evaluated 1-week post-training. After training, the groups did not differ in objective skill level, but drivers who completed the skills training rated their skills significantly higher and were more likely to overestimate their skid handling ability. This highlights the need to carefully structure vehicle skills training so that it does not encourage novice drivers to overestimate their abilities.

3.3.4 Sweden: Cooperative professional–private pre-licence training

Another evaluation of experimental driver training in Sweden involved a programme of systematic cooperation between professional driving instructors and lay supervisors of pre-licensed drivers (Gregersen 1994). The control group included 947 learners who received only private training (ie, learning to drive with a lay supervisor such as a parent). The training group were 846 learners who also received 16 professional lessons (~45 minutes each), of which their private supervisor had to attend at least 4 lessons. For part of each lesson, the learner practised 'commentary training' in which they were asked to report actions, observations, perceived risks, and potential preventive measures. The learner and supervisor were given instructional training materials to guide their practice throughout the learning period. Finally, the learner was given a series of exercises to reduce overconfidence in their driving skills.

All participants in Gregersen's (1994) evaluation were 17 years old at the time of training and were followed up for their first 2 years of licensure. In the first year, self-reported crash rates were significantly higher for the experimental training group than the comparison group (0.44 vs 0.30 crashes per 10,000 km). In the second year of licensure the pattern reversed, with the comparison group being significantly higher than the training group (0.24 vs 0.17 crashes per 10,000 km). The reason for this unexpected pattern of results is unclear. Process evaluation of the training programme indicated that not all professional instructors implemented the programme as intended; however, inadequate implementation would be more likely to produce a neutral effect in both years, rather than the observed negative impact in year 1 that shifted to a positive impact in year 2.

Gregersen (1994) hypothesised that perhaps the skills targeted may have been too challenging for the initial stage or skill level of the novice drivers. That is, because they were still learning basic procedural skills handling, trying to teach them additional skills such as commentary training may have cognitively overloaded them, resulting in reduced safety in their first year of unsupervised driving. Then, once they mastered procedural skills, they were able to implement the cognitive skills targeted by the commentary training. This is an interesting hypothesis, although there are no data to support it and it would be difficult to test empirically. Nevertheless, it highlights the need to introduce new skills progressively at a stage that is appropriate for the learner's development.

3.3.5 Netherlands: 'Second phase' training

De Craen et al (2005) evaluated a 'second phase' training course for young drivers, which involved test-track training and group discussions. Unfortunately, several practical constraints interfered with planned evaluation, which was run at two sites. Both the training and the control participants were supposed to complete two instructor-assessed on-road drives with feedback: one before training and one a month later. One of the sites experienced extreme weather during the first assessment drive, which appeared to inflate scores; the authors speculated that instructors were impressed with how the novice drivers handled the adverse conditions. This resulted in several statistical analyses yielding a significant interaction between measurement time (pre- vs post-training) and location, with one site showing a significant improvement and the other showing a significant decline in scores relating to driving skills and calibration.

A further issue with the evaluation was participant recruitment and retention, with most potential participants either declining to participate or failing to complete all sessions. This resulted in a substantially smaller sample size than planned (500 were approached; 376 signed up; only 127 completed). De Craen et al (2005) suggested this indicates most young drivers are not interested in voluntary post-licence training. Overall, the sampling bias, combined with the unusual weather conditions at one of the two sites, means that the conclusions from this study are not generalisable.

3.3.6 Spain: 1-day ‘second phase’ training

Molina et al (2007) evaluated a ‘second phase’ post-licence novice driver training course developed by the Real Automobil Club de Cataluña (RACC) in Spain. The 1-day course involved three 90-minute sessions, which could be completed in any order: a test-track drive, an on-road drive, and a group discussion. The course included elements of vehicle handling, including skid training, but was designed to emphasise insight training or self-awareness of limitations as well as issues such as resisting peer pressure and avoiding showing off.

The evaluation included 238 young drivers (aged 18–24) with less than 3 years’ driving experience, who were randomly assigned to either training or control and were surveyed pre-training and 9-months post-training. Self-reported careful driving skills significantly increased for the training group, but not for the control group. There were no other significant differences between groups in terms of self-reported driving habits (speeding, driving in an improper state) or perceived risk. The authors noted their sample was not representative of the broader young driver population due to non-response, non-random dropout, and the fact that all were RACC members.

3.3.7 Montana, USA: 1-day defensive driving

In 2005 the Department of Transportation in Montana, USA, designed and evaluated a 1-day defensive driving course, which was targeted at young novice drivers who had previously completed basic driver education. Stanley and Mueller (2010) evaluated the results over a 4-year period (2006–09), with participants randomly allocated to either training or control group. The study began with 347 participants, but by the final survey only 50% of the trained group and 47% of the control group remained.

The analyses revealed no significant association between training and official crash rates or official citations. Some self-report measures showed significant differences between groups, with trained drivers being significantly *less* likely than controls to report citations in year 1 (20% vs 30%) and year 2 (31% vs 45%), but significantly *more* likely to report single-vehicle crashes in year 3 post-training. After controlling for driving exposure, these differences were no longer statistically significant. Stanley and Mueller (2010) suggested that participants may have misremembered details of their citations, or misunderstood the definition of a citation, which could explain the discrepancies between self-reported and official data.

Although participating in the training programme did not have a clear effect on safety, the authors found that ‘report card’ assessments during training correlated with later violations and crashes.

This study contained some strong design elements, as it used random assignment to training condition and long-term follow-ups. However, the sample size is too small to detect changes in crash rates, and there are some obvious statistical errors (incorrect numbers in the report) which undermine confidence in the findings.

3.3.8 Belgium: ‘On the Road’

Brijs et al (2014) evaluated ‘On the Road’, a short post-licence programme designed to target higher levels within the Goals of Driver Education (such as insight). The programme included 3.5 hours of training, which involved classroom, test-track and on-road sessions, plus 3 weeks of personalised weekly feedback emails from an instructor. Participants were 366 young novice drivers with an average of 14 months’ solo driving experience pre-training, who were randomly assigned to training or control groups. All had voluntarily signed up for training, so the control group was assessed before they completed training. The evaluation focused on risk knowledge, risk perception, behavioural intentions, self-efficacy (ie, controllability of actions) and norms (ie, what others do/think I should do, what I think I should do).

The training intervention significantly increased risk knowledge scores: this effect size was large and maintained at follow-up 2 months later, but the level of risk knowledge was still relatively low. Immediately post-training there was a significant positive effect on descriptive norms, behavioural intentions and self-efficacy relating to speeding, but all effect sizes were small and the effects were not maintained at follow-up. Conversely, training appeared to have a negative effect on norms and risk perception relating to drink driving, although again these effects were small to moderate in magnitude and were not maintained at follow-up. The authors concluded that the intervention only had a meaningful impact on knowledge, not attitudes or behaviour, but noted that all participants were highly safety conscious before the intervention, so this could reflect a ceiling effect whereby it is difficult to observe further improvement in these measures.

3.3.9 Italy: Advanced driver training

In Italy, advanced driver training courses are endorsed by the Ministry of Transport and Infrastructure if they meet established standards and criteria. Usami et al (2016) evaluated approved advanced driver training courses that are conducted throughout Italy. The courses involve 2 hours of theoretical instruction and 5 hours of practical instruction, focusing on vehicle control and mastering traffic situations.

Usami et al’s (2016) main evaluation compared 6,932 drivers who completed the training course in 2011–12 with a control group of 8,469 drivers who did not undertake a course during the same period. Violations decreased significantly for the training group but not the control group. When examining subtypes of violations, they found speeding violations, which accounted for half of all violations, *increased* slightly for both treatment and control groups, whereas other types of violations reduced after training (seatbelt, phone use, signal violations). Before–after comparison of drivers who attended training indicated crashes reduced post-training, but there were no comparable data for a control group, so this trend is hard to interpret.

Notably, the training and control groups had pre-existing differences in violation rates, as well as demographic characteristics. The treatment group were predominantly males aged 22–44, whereas the control group had a more equal gender balance and wider age distribution. When analysing a subsample of drivers matched for violation rates and key characteristics, their analysis showed that violation rates (especially speeding) increased after training but declined for the control group over the same period. This highlights that even when training has an overall neutral or positive effect when measured at the population level, it can have a negative effect on some drivers. Also, similar to the DeKalb study, it highlights that different analyses of the same dataset can reach opposing conclusions.

3.3.10 Queensland, Australia: Holden 'driver improvement' course

Petersen and colleagues evaluated a 2-day advanced driver training course that was conducted at the Holden Driving Performance Centre in Queensland, Australia (Petersen and Barrett 2009; Petersen et al 2006, 2008). The training programme was a 'driver improvement' course for experienced drivers, which was designed to teach principles of safe driving and introduce participants to a specific driving technique (ie, physical position and approach to manoeuvring the vehicle). All participants attended two driving tests, several days apart. Participants in the training condition completed the 2-day training course between the two driving tests, whereas control participants only did the test sessions.

Data are reported in three separate papers, which report slightly different sample sizes of 21–26 participants for the training group and 12–13 for the control group, but identical values for mean age, standard deviation of age, mean years of driving experience and standard deviation of driving experience. This suggests the three papers are reporting different outcome measures from the same study, but it is unclear. One possibility is that it is the same sample, so the authors reported the original/full sample characteristics, even though some individuals were not included in all papers (eg, because of missing data). This is problematic because if the three papers are evaluating unique administrations of the course, that would provide stronger evidence than if all three were evaluating the same groups (the latter creating a false impression that the course has been repeatedly shown to improve behaviour).

With that caveat in mind, the evaluations indicated that the training programme was associated with improved postural stability, which in turn influences vehicle manoeuvring and control when cornering and changing lanes (Petersen and Barrett 2009; Petersen et al 2008). Training also facilitated smoother braking performance, but when driving a car with an anti-lock braking system, trained drivers took longer to stop completely than control drivers (Petersen et al 2006). These results show that it is possible to significantly improve vehicle handling skills.

Although the Holden training programme is relatively advanced, approximately half the trainees were on their provisional (restricted) licence. However, it is unclear whether the benefits observed would translate into meaningful safety benefits on the road, such as reduced crashes. It is also unclear how participants were recruited, but it appears they self-selected into training, which means the results may not generalise to a broader population of drivers who are less motivated to seek training.

3.3.11 UK: Institute of Advanced Motorists coaching programme

Stanton et al (2007) evaluated an advanced driver coaching programme run by the Institute of Advanced Motorists in the UK. The programme used the 'IPSGA' hazard model: Information, Position, Speed, Gear and Acceleration. Drivers are coached to constantly seek information and position themselves so they can safely negotiate any hazards. When they encounter a hazard, they should adjust their speed, gear (for manual vehicles) and acceleration appropriately so they remain safe and drive smoothly. The evaluation included 25 experienced drivers who completed the coaching programme, compared with two control groups of 25 drivers each: one active control group completed accompanied drives with no coaching, and one passive control group completed only the evaluation activities, which were conducted 8 weeks apart.

The programme was evaluated in terms of its impact on 'knowledge, skills and attitudes'. Skills were measured as observable driving behaviour, with the trained group showing significant improvements in speed, steering, headway, mirror use and gear changes, whereas the control groups showed no change. Knowledge or situation awareness was measured by having participants complete a post-drive booklet about critical situations they encountered. After training, participants' situation awareness was more complex, including an increased amount of information, new information concepts, and greater connection between concepts. The control groups showed no change. Attitudes were measured by examining locus of control (ie, internal/external attributions for behaviour), which also improved slightly for trained drivers but not controls.

3.3.12 New Zealand: Vehicle handling vs higher-order skills

Isler et al (2011) conducted a randomised controlled trial with 36 teenage drivers divided into three groups: vehicle handling skills, higher-order skills, or control. The training programmes lasted 1 week each and were conducted at a 2-week camp. The evaluation used a time-lagged design, which meant that after the initial training and evaluation, participants then received the alternative training; therefore, the long-term effects of the two programmes could not be evaluated.

The vehicle skills course involved both on-road and track driving and focused on vehicle manoeuvring, including parallel parking and steering around curves. Participants observed each other so they could give and receive peer feedback on vehicle handling skills. The higher-order cognitive skills programme was designed to improve hazard perception, situation awareness and self-assessment of skills. Both groups were evaluated on vehicle handling, hazard perception and attitudes. Both groups improved on the skills targeted by training; that is, the vehicle skills group improved their on-road skills but not their hazard perception skills, attitudes, or confidence. In contrast, the higher-order skills group improved on metrics associated with safety, including improved hazard perception, visual search and attitudes, and reduced driving confidence. Although this is only a small, short-term evaluation, it highlights the importance of targeting relevant skills, and the fact that ‘better’ skilled drivers are not necessarily safer.

3.4 Simulator training

Driving simulators offer a potential tool for driver training. An advantage of simulator training is that learners can be safely and efficiently exposed to a diverse range of situations, including critical situations that could result in a crash.

Simulator training has been used in a variety of contexts, as described in the following subsections. Broadly, these can be divided into programmes that are designed to teach basic procedural skills to learner drivers and programmes for licensed drivers that target specific skills or scenarios.

3.4.1 Simulator training for learner drivers

Several studies have evaluated simulator training for learner drivers; however, these evaluations differ substantially in nature and scope. Some programmes are experimental or pilot programmes, evaluated by researchers (and often designed by those same researchers). Other evaluations report real-world programmes; for example, where learner drivers can undertake simulator training in a driving school.

3.4.1.1 Experimental evaluations

The review identified three papers reporting experimental evaluations of simulator training for learner drivers. All were conducted with teen drivers in the USA.

Cox et al (2009) reported a pilot study evaluating simulator training for teenagers with no driving experience. The evaluation included 20 participants (10 trained, 10 control) who completed a short expert-assessed on-road drive. A ‘high-fidelity’ simulator (ie, relatively realistic) was used, with realistic vehicle controls and 180° visual field of view displayed. The simulator was used to provide scaffolded training, starting with no traffic and becoming increasingly complex. Trainees were instructed to focus on maintaining lateral position and speed, braking smoothly, and using turn signals and mirrors. Trained participants were rated as significantly better than control participants on most outcome variables, which suggests that simulator training can be useful for initial procedural skills training. However, a major limitation is that none of the participants had driven previously and the control group did not receive any alternative training, so training was confounded with overall driving experience and knowledge. The groups also differed significantly in age, with the control group being older (mean age 18.2 vs 16 years). Although older drivers are typically safer, in jurisdictions that allow 16-year-olds to drive, there are likely to be systematic differences among those who reach 18 without having obtained a licence (eg, less interest in driving, or greater anxiety about driving), which could also have influenced the results.

Another short-term small-scale evaluation of simulator training was undertaken by Allen et al (2011) with 67 driver education students from a single high school. The authors did not report how many participants received training, or how they were allocated to training versus comparison groups. The training involved eight short (10–20 minute) sessions taken over several weeks as part of regular driver education. A 'low-fidelity' simulator was used, which comprised a single monitor and gaming controls. This is relevant because it is a low-cost setup that could be implemented at multiple locations. The trained group showed significantly greater improvement than the control group on nearly all measures, with superior collision avoidance and fewer speed exceedances and intersection violations. These differences could not be attributed to comfort or skill at manoeuvring the simulator. The major limitation of this study is that it involves a small sample, with short-term follow-up only, and the outcome was simulated driving rather than real-world driving performance.

To provide a more extensive evaluation of simulator training, Allen et al (2012) conducted a longitudinal evaluation with 554 teenagers, comparing three types of simulators:

1. a low-fidelity single-monitor desktop simulator with gaming controls, representing a 45° field of view, similar to Allen et al (2011)
2. a three-monitor desktop simulator with gaming controls, representing a 135° field of view
3. a high-fidelity simulator incorporating a full vehicle with an instrumented cab and the road images projected on a screen representing a 135° field of view.

Participants were allocated to one of three simulator types, based on their recruitment location. The single-monitor simulators were situated in high schools with participants recruited from driver education classes, whereas the other two simulators were in research labs with participants recruited through local Department of Motor Vehicles offices. The same training programme was used in all simulators, consisting of an orientation slideshow and a series of 15-minute training runs. Participants completed 6–9 training runs, depending on how they performed (ie, some needed longer to pass the 'exit test' assessment).

Official crash records for the first year of licensed driving were compared with population records of novice drivers who had not been part of the evaluation (Allen et al 2012). This analysis revealed different patterns for the single-monitor simulator compared with the three-monitor and vehicle cab simulator, suggesting using a simulator with a wide field of view is crucial for training. Specifically, participants who were trained on wide field of view simulators had significantly lower crash rates (6%) compared with both untrained participants and those trained on the narrow field of view (single monitor) simulator (13% crashed). Crash rates were not significantly different when comparing untrained participants and the single-monitor training group. This is consistent with data collected during training, which suggested that participants drove differently when using the single-monitor simulators compared with other types of simulators (eg, treating it more like a computer game).

Overall, Allen et al's (2012) longitudinal evaluation suggests that simulator training is less effective if it uses low-fidelity simulators. However, a substantial limitation of their research is the non-random group assignment, which means there are likely systematic confounds between participants allocated to different simulators (including age and location). The study also lacked a proper control group comparison, which was partially addressed by analysing population crash data; however, there are likely systematic differences between the sample who volunteered for simulator training and the broader population, meaning the groups are not equivalent. What is unknown is whether eliminating these confounds would change the conclusions of the study.

3.4.1.2 Real-world evaluations

In some countries, learner drivers can complete simulator training as part of their broader programme of driver training (ie, at a driving school). The review identified two studies evaluating this type of training; obviously, both involve participants self-selecting into the simulator training group, and therefore the training condition is potentially confounded with other factors (eg, interest, motivation, geographical location, finances).

Rosenbloom and Eldror (2014) compared drivers who had completed 4–6 hours of pre-licence simulator training at an Israeli driving school with drivers who did no simulator training. All had been licensed in the preceding 12 months. Their evaluation included a survey study with 280 drivers (140 simulator trained, 140 control) and an expert-assessed on-road drive with a subset of 40 participants (20 simulator trained, 20 control). The survey assessed self-reported attitudes, intended behaviours and norms. Most measures were not significantly different when comparing simulator-trained drivers with controls, with the only significant differences being small in magnitude. Specifically, simulator-trained drivers showed more brake pedal presses and fewer headway events, suggesting that they were more actively monitoring headway, but they also reported lower safe driving intentions. The authors concluded that there was no specific advantage or disadvantage to simulator training. However, it is worth noting that their sample is relatively small, especially for the on-road component.

Hirsch and Bellavance (2017) conducted a much larger evaluation of simulator training in Quebec, Canada, comparing 1,120 simulator-trained drivers, who volunteered for the research, with the broader population of over 180,000 novice drivers who had not completed training. Learners could substitute simulator training for on-road training for up to 6 of the mandatory 15 hours. The scenarios focused on ‘vision training’, although details are not provided. Teachers used their own judgement to determine how many simulator lessons to use for each student and which scenarios to include. Nearly all students in the trained group (95%) completed ≤ 4 hours of training, with 28% completing only 1 hour.

Participants were evaluated for their first 2 years of unsupervised driving, with analyses adjusted for age and number of days owning a vehicle (which was used as an indirect measure of driving exposure). Rates of officially recorded crashes were not significantly different between trained and untrained drivers. Simulator-trained drivers were less likely than the comparison group to commit infractions (violations), but differences were only statistically significant for females in the period 18–24 months post-licensure, and for males during the periods 6–24 months post-licensure. This could represent a delayed effect of training (which would be unusual, as it would be more common that training effects would decay over time). Alternatively, it could reflect the fact that training has a relatively modest effect on violations, which is not always significant.

3.4.2 Simulator training for licensed drivers

Simulator training for licensed drivers can be divided into two broad categories: those using standardised programmes that are the same for all drivers, and those that provide individualised training and/or feedback based on the trainee’s performance.

3.4.2.1 Standardised programmes

The review identified three papers evaluating ‘standardised’ simulator training programmes for licensed drivers. These programmes target specific skills or situations, but all trainees receive the same programme.

All involve experimental programmes, and the evaluations are unfortunately limited by very small sample size, which limits the generalisability of any conclusions and the overall usefulness of the evaluations.

Ekeh et al (2013) evaluated a simulator training package that comprised 12 short modules that were designed to progressively develop visual search and vehicle control skills while driving. Each module contained an assessment drive, on which the trainee needed to score 70% to progress. The total training time was 5.5–6 hours and all modules could be completed within 1 week. The evaluation was described as a 'pilot' and included 35 participants randomly allocated to training or control. Participant were recently licensed teen drivers, whose official crash and violation records were monitored for 18 months post-training. There were no statistically significant differences between groups, but the sample size is so small the analyses are underpowered. For example, at the 18-month follow-up, 1/16 trained drivers (6.2%) and 4/19 control drivers (21.1%) were involved in a crash. The relative percentages show a positive trend, but the numbers are far too small to draw meaningful conclusions. The authors' main conclusion was that simulator training for novice drivers was feasible and should be explored further with larger samples.

Zhao et al (2019) evaluated simulator 'perception–norm–execution' training designed to reduce risky driving behaviours in experienced drivers. In the 'perception' phase they completed a simulated drive with critical situations, in which they would crash if they engaged in risky behaviour (eg, unsafe lane changes or overtaking, speeding, running red lights, distracted driving). In the 'norm' phase they were shown videos of risky crash-causing behaviour, which was explicitly contrasted with correct safe behaviour (ie, what they 'should' do). Finally, in the 'execution' stage they were guided through and asked to reflect on risky simulated driving situations. The evaluation included 44 experienced drivers (average 10 years' driving experience), with half randomly assigned to receive the simulator training and half completing video-based training. Their simulated driving behaviour was compared before, immediately post- and 3 days post-training. Most analyses focused on before–after comparisons, and both programmes showed initial positive effects (eg, improved signal use, reduced speeding, safer deceleration), but the effects were larger for the simulator group. Three days post-training, the simulator group still showed some positive effects, albeit less pronounced. This study has some notable limitations: aside from the small sample size and short-term follow-up, there are some inconsistencies and omissions in the paper, and not all statistical tests were fully reported, which makes it hard to assess its quality.

In an example of a highly specific simulator training programme, van Leeuwen et al (2013) evaluated simulator training designed to encourage young novice drivers to distribute their gaze more effectively. Participants were 30 young drivers with an average age of 19.1 years and just over 6 months' solo driving experience, who were split into two approximately equally matched groups. Both groups completed five short simulator drives while wearing an eye-tracker. In the training condition, participants had to respond to purple dots that appeared randomly on the road, by the left and right roadside, and near the rear-view mirror; the purpose was to encourage them to distribute their gaze more to key areas. Control participants simply drove the simulator. Trained participants showed greater gaze distribution during training, but the effects were not maintained in the transfer drive when the purple dots were disabled, suggesting that no actual learning or skill development occurred.

3.4.2.2 Simulator training with individualised feedback

Some researchers have exploited simulator capabilities to deliver individualised training, or training with individualised feedback, to drivers based on their performance.

Biondi et al (2020) reported a brief experimental evaluation of precision teaching for lane maintenance, which involved giving drivers feedback when they deviated from optimal lane position during a simulated drive. Participants were drivers aged 20–29 with at least 1 year of experience; 29 received auditory feedback only; 19 received auditory feedback plus redundant visual feedback; and 21 received no training. Participants in the feedback conditions were more likely to maintain optimal lateral position and had reduced standard deviation of lateral position, compared with those who received no feedback. There was no difference in the extent of improvement when comparing auditory-only with auditory and visual feedback. The limitations of this study are its small, unrepresentative sample size, and the fact that the training was administered and evaluated within the same session. It is unclear whether these effects would be sustained long-term, whether they would generalise to real-world driving, and if so whether they would meaningfully improve safety. This paper is also missing key details about the training and evaluation process, including the data collection location (which is noted as 'University of XX').

Ka et al (2020) developed simulator training to target multiple types of risky driving: speeding, reckless speed changes, rapid acceleration/deceleration, erratic steering control, erratic lane changes, and tailgating. The training programme used surrogate safety measures of these behaviours (eg, automatically calculated values representing dangerous behaviour) to give tailored feedback to drivers. The evaluation used a before–after design with no control group and included 21 novice drivers with less than 2 years' experience (as well as 21 commercial drivers and 16 older drivers). The novice drivers significantly improved on nearly all safety measures from pre-training to post-training. However, the lack of a 'no training' control group means that training is confounded with general simulated driving practice. To properly evaluate this type of programme, the researchers would need to include a comparison group that spent an equivalent amount of time driving in the simulator without personalised feedback and see whether their behaviour changed with practice.

Lang et al (2018) compared personalised training within a virtual reality simulator (ie, wearable goggles to display the road environment) with four alternatives: standard training along a pre-set simulated route, watching a training video, reading a training manual, and no training. The personalised training triggered specific scenarios based on the undesirable habits that drivers exhibited (eg, failing to check mirrors or use signals, not stopping for pedestrians or slowing past cross-roads). Participants were randomly assigned to training condition, with 10 drivers per group. Immediately post-training, the personalised training group showed the most 'good habits' and had a faster response time when braking for a pedestrian, and these positive effects were maintained at follow-up a week later. The standard simulation and video training groups also showed small benefits, but performance for these groups dropped at follow-up. Although this study implies that personalised training is beneficial, the evaluation is severely limited. The sample size is extremely small, and the authors do not provide demographic or driving information about the participants other than a range of '1–20 years' of driving experience. As with Ka et al (2020), training is confounded with simulation experience, so it is hard to assess which benefits are specific to the training and which are attributable to general practice using the virtual reality simulator.

3.5 Hazard perception

Almost half the papers identified in the systematic review focused on hazard perception interventions (and, of course, many of the other types of training programmes include elements of hazard perception too). Although there is a lot of research on hazard perception, including hazard perception training, much of it arises from the same research group at the University of Massachusetts Amherst, who have produced, designed, and evaluated several training programmes.

Most evaluations of hazard perception training are experimental and have similar strengths and weaknesses. In most cases, participants are randomly allocated to training versus control or comparison groups. A drawback is that the sample sizes are usually small, with several studies having 10 or fewer participants per group, which means that even with random assignment there may be systematic differences between groups. Also, small samples are usually not representative of the broader population because they represent a minority of individuals who were willing to volunteer for training – and participate in the research. Finally, most of the evaluations examine only short-term effects and do not assess real-world driving behaviour (instead using simulator driving or hazard perception test score as the outcome measure) so it is unclear whether any benefits of training would translate to meaningful improvements in on-road safety.

Because there are so many evaluations of hazard perception training, they have been grouped according to intervention characteristics. Some programmes have been evaluated multiple times, or used as the basis for subsequent training programmes, and therefore have their own subheading below. These include Driver ZED, DriveSmart, Risk and Awareness Perception Training (RAPT), Engaged Driver Training System (EDTS), Forward Concentration and Attention Learning (FOCAL) and Anticipate–Control–Terminate (ACT) training. These programmes are similar in that all involve video, computer or tablet-based training designed to help drivers anticipate hazards or avoid errors. The remaining programmes have been grouped according to the broad type of training: verbal commentary, general hazard anticipation, and specific hazard perception (ie, training for a specific type of hazard, such as pedestrians or motorcycles, or to perceive hazards in a specific context, such as driving on rural roads).

3.5.1 Driver ZED

Driver ZED (Zero Errors Driving) was an early PC-based training programme produced by the American Automobile Association (AAA) Foundation for Traffic Safety in the USA. The programme was designed to raise risk awareness and involved watching 80 driving videos. After each video, the driver was asked to answer questions about the video, click on potential hazards and identify potential actions and when they should be taken.

Fisher et al (1998, 2002) randomly allocated young novice drivers aged 16–17 to either Driver ZED training or control, and then evaluated them in a driving simulator. Both studies involved very small sample sizes with no long-term follow-up. They also selectively reported only 2–6 of the 24 assessment scenarios. Fisher et al (2002) state they selected the scenarios designed to best test the training, but it is unclear whether this selection occurred before or after the results, and whether positive effects occurred throughout the drive or only for the scenarios selected.

Neither of the two scenarios reported by Fisher et al (1998) showed significant differences in speed between trained and untrained drivers, but these analyses were underpowered, with only 8 drivers per group. The subsequent evaluation by Fisher et al (2002), with 15 drivers per group, found that trained novice drivers were significantly better than untrained novices on several measures, including reaction to latent/hidden hazards and potential hazard warnings (eg, anticipatory braking). Novices were also compared with a group of untrained but more experienced drivers (who were 5–10 years older than the novices). Experienced drivers exhibited safer behaviour than novices, but the trained novices appeared to become more like experienced drivers in their risk anticipation and less like untrained novices.

3.5.2 DriveSmart

Another early hazard perception training programme was DriveSmart, a CD-ROM training package developed by Monash University Accident Research Centre in Melbourne, Australia. This programme focused on attentional control, training drivers to prioritise specific tasks, switch or share attention between tasks, and better understand their abilities and limitations while driving. The programme was designed to complement supervised practice for learners by targeting specific risky driving situations.

Regan et al (2000) evaluated the simulated driving performance of 103 learner drivers, who were randomly assigned to training or control. The authors reported that the trained group showed a ‘positive’ effect in 15 out of 32 scenarios assessed and neutral effects in the others. The training group also displayed better attentional control, in that they adjusted faster when the posted speed limit changed and drove closer to the speed limit than control drivers (who had significantly faster mean speeds).

An earlier, separate evaluation of the attentional control training component alone (Regan et al 1998) with 72 learner and probationary (restricted) drivers found that the only significant effect on driving performance was that trained drivers exhibited significantly higher maximum acceleration (suggesting greater effort required to adjust to a new higher speed zone). Regan et al (1998) suggested that perhaps the driving headway task used during training was too difficult. This highlights the importance of piloting training tasks and separately testing individual elements of multi-component training programmes to better understand which components are likely to be effective.

3.5.3 Risk Awareness and Perception Training (RAPT)

The single most evaluated hazard perception training programme is the Risk and Awareness Perception Training (RAPT) PC-based programme, which has had several iterations and variants. It was developed by researchers at the University of Massachusetts Amherst in the USA, although versions of RAPT have been adapted and evaluated by other researchers, including in the UK.

In the first two versions, RAPT-1 and RAPT-2, participants watched bird’s eye (overhead plan) views of developing traffic scenarios, which were designed to encourage them to better understand spatial relations between hazards. For example, the scenario might involve a vehicle approaching an intersection. The task required them to place circles on any locations the driver needed to monitor, and on any locations that could contain concealed pedestrians or vehicles that could pose a future hazard. After making their response, the participant was given feedback showing correctly marked locations and a marker representing the driver’s field of vision to highlight areas of their view that were obstructed (Pollatsek et al 2006). The only difference between RAPT-1 and RAPT-2 was that RAPT-2 included additional scenarios, based on participants’ weaknesses identified in earlier research.

Evaluations of RAPT-1 and RAPT-2 in a driving simulator with high school student learner drivers found that RAPT-trained drivers were significantly more likely to fixate on areas containing potential hazards (Fisher et al 2006). Specifically, trained drivers fixated on the prospective hazard locations 52–57% of the time, whereas controls fixated on these areas only 28–35% of the time. Both evaluations used small samples (48 and 24 drivers), randomly allocated to training or control, with effects evaluated immediately or shortly after training (3–5 days later).

In RAPT-3, the programme was modified to provide the driver’s perspective view. This programme was initially evaluated with small-scale studies (24 drivers, randomly allocated to training or control) of novice drivers (1–4 years’ driving experience) using both simulated and real driving. From the details reported in individual papers, it is unclear whether the authors conducted multiple field studies evaluating RAPT-3, or if they conducted a single field study and reported it multiple times (Fisher et al 2006, 2007; Pradhan et al 2009). Similar results are reported across studies, with RAPT participants being significantly more likely to glance at critical regions that could contain or conceal potential hazards. The training showed both near-transfer (ie, benefits for scenarios highly similar to training) and far transfer (ie, benefits for more distinctive scenarios), although there were some critical scenarios for which trained participants did not perform better than controls (Pradhan et al 2009).

Although experimental evaluations of RAPT have consistently shown significant positive changes in glance behaviour, indicating hazard perception, studies that have more directly assessed driver safety have not been as conclusive (McDonald et al 2015; Thomas et al 2016).

In a small experimental simulator evaluation, McDonald et al (2015) found no significant differences in performance between trained and untrained drivers on a left-turn task (equivalent to right turn across traffic in New Zealand).

In a large-scale evaluation of 5,251 recently licensed drivers aged 16–18, Thomas et al (2016) found that RAPT training had no significant effect on female drivers. There was a significant effect of training for males, with RAPT-trained participants experiencing approximately 24% lower crash rates than controls in their first 12 months of licensure. These gender differences were unexpected because no previous research had indicated gender differences in hazard perception training.

More recently, some researchers have adapted RAPT-3 for use in a wearable head-mounted virtual reality display, which provides a more immersive experience of the scenarios. Evaluations have indicated that training using head-mounted displays is even more effective than PC-based versions of RAPT, but these evaluations only assessed effects immediately post-training and again used very small samples, with 12–19 participants per condition (Agrawal et al 2018; Madigan and Romano 2020).

Some other hazard perception training programmes have different names but appear to be closely inspired by RAPT. Krishnan et al (2019) evaluated the Secondary Task Regulatory and Anticipatory Program (STRAP), which used the same format as RAPT-1 and RAPT-2, except it was run on Microsoft PowerPoint. Samuel et al (2013) evaluated Road Aware, a simulator training programme modelled on RAPT that was designed to provide greater diversity in training (presenting multiple perspectives, and multiple trials of each type of scenario). Both evaluations included 48 young, licensed drivers (equivalent to early full licensure in New Zealand), randomly allocated to either training or control. When tested in a driving simulator, in both studies trained participants were more likely to make anticipatory glances towards latent hazard locations. STRAP also evaluated the impact of a secondary task and found that trained drivers had a longer ‘heads up’ time (ie, they were better focused on the primary task of driving with less attention to the secondary task).

3.5.4 Engaged Driver Training System (EDTS)

The Engaged Driver Training System (EDTS) is a tablet-based hazard perception training programme for novice drivers, which takes approximately 30 minutes and is based on RAPT and RoadAware. EDTS has participants ‘steer’ through scenarios by rotating the tablet, with acceleration and braking controlled by on-screen buttons. The programme provides error feedback and is especially targeted towards identifying latent hazards and reducing distraction.

EDTS has been evaluated in separate studies to test its effects immediately post-training (Krishnan et al 2015), 1-week post-training (Ahmadi et al 2018; Zafian et al 2016) and 6 months post-training (Ahmadi et al 2018). All evaluations used small samples (20–35 drivers total, split between training and control).

The first evaluation, immediately post-training, used simulated driving and found that EDTS significantly improved latent hazard detection and hazard ‘clue’ detection (Krishnan et al 2015). The trained group also displayed more ‘heads up’ time during a designated secondary task (ie, better focus maintained on the primary task of driving).

Subsequent evaluations (using different participants) examined performance during a real drive on local public roads. Analysis of eye movements 1-week post-training found EDTS-trained drivers were significantly more likely to glance at latent hazard locations (71% vs 27% of the control group; Zafian et al 2016). A subsequent evaluation combined eye movement data with verbal commentary made by participants during the drive to assess their situation awareness, and specifically their ability to project future situations. At 1-week post-training, EDTS drivers exhibited significantly better situation awareness at crosswalks, four-way intersections and while turning, but not at roundabouts. Six months later, both groups showed reduced situation awareness, and the trained group was only significantly better for crosswalks (Ahmadi et al 2018).

This highlights the importance of conducting longer-term follow-up evaluations as impressive initial effects may decay relatively quickly.

3.5.5 Forward Concentration and Attention Learning (FOCAL)

The Forward Concentration and Attention Learning (FOCAL) programme is a computer-based attentional training programme which uses the ‘3M’ error approach: make mistakes, mitigate mistakes, master skills. During training, participants must locate a street on a map while viewing video clips, which teaches them to task switch between navigation and watching the road. Specifically, they are taught to limit ‘off-road’ glances to less than 2 seconds.

Pradhan et al (2011) reported an initial small-scale evaluation of FOCAL using a sample of 37 novice drivers (learner or restricted licence stage) aged 16–17, who were randomly allocated to training or control. Participants were evaluated in a real drive on public roads. FOCAL-trained participants made significantly fewer off-road glances that were longer than 2 seconds, and most of the other glance measures also indicated significantly better performance for trained drivers.

Divekar et al (2016) evaluated FOCAL in a sample of 40 parent–teen dyads (80 participants total), randomly assigned to training or control. Teens were aged 16–18 and were learner drivers, whereas their parents had at least 10 years’ driving experience. Participants were tested in a driving simulator immediately post-training and again after 4 months. Immediately post-training, there was a significant effect of experience (parents made fewer long glances than teens), but the effect of training was not significant, although there was a trend towards FOCAL-trained drivers making fewer long glances than untrained novices (25% vs 33% of glances > 2 seconds). At 4-month follow-up, there was a significant positive effect of training on glance behaviour, with FOCAL-trained novices exhibiting significantly fewer long glances (19% vs 29%). In contrast, FOCAL-trained experienced drivers showed worse performance at follow-up compared with immediately after training. The reason for this is unclear.

A further evaluation by Unverricht et al (2020) found similar benefits of FOCAL training in a slightly older novice driver sample comprising 36 university undergraduates aged 18–21 with just over 2 years’ driving experience. Trained drivers made fewer long off-road glances than controls (20% vs 36% of glances > 2 seconds) and had a better calibrated assessment of their skills. Specifically, the control drivers overestimated their ability to limit long glances, whereas the trained drivers did not. This suggests that FOCAL teaches drivers to control and estimate their glance duration effectively (ie, that it improves the specific driving behaviour targeted by the programme).

3.5.6 Anticipate–Control–Terminate (ACT) training

Anticipate–Control–Terminate (ACT) training is a programme designed to enhance hazard mitigation, designed by the same team responsible for RAPT and FOCAL. As with FOCAL, it uses error-based training. First participants complete a driving game in which they are shown a scenario and must indicate how they would act (speed, steering, where they would look). Scenarios are designed to represent fatal crash risks. Participants are then shown choices made in that scenario by experienced drivers, who either had or had not crashed in the preceding 10 years, to show them appropriate and inappropriate strategies for dealing with the hazard.

Muttart et al (2019) evaluated ACT with a sample of 36 newly licensed teenage drivers, who were randomly allocated to training or control. ACT-trained participants completed pre- and post-test simulator drives, whereas control participants only completed a post-test drive. Trained drivers were significantly less likely to crash than controls, and they exhibited more slowing behaviours and anticipatory glances.

The ACT programme is newer than its ‘sister’ programmes RAPT and FOCAL, so obviously further testing is required to confirm whether training benefits persist over time and transfer to real-world driving.

3.5.7 Combined programmes: ACCEL and SAFE-T

RAPT, FOCAL and ACT are designed to specifically target hazard anticipation, attentional maintenance, and hazard mitigation, respectively. In addition to evaluating these programmes separately, the researchers have created combined programmes that include elements of two or all three programmes.

3.5.7.1 ACCEL

The Accelerated Curriculum to Create Effective Learning (ACCEL) programme combines elements of RAPT and ACT. It was designed as an open-access programme that could be downloaded from the internet to any computer, laptop, tablet, or smartphone. The training takes approximately 2 hours. Fisher et al (2017) evaluated ACCEL, comparing three groups: 50 ACCEL-trained novice teenage drivers, 25 untrained novice teenage drivers and 25 experienced drivers. Participants were assessed in simulator drives before and after training, with post-training follow-ups immediately after training and between 3 and 6 months later, although only 26 of the novice drivers returned for this stage.

Immediately after training, ACCEL-trained participants were significantly better than untrained novices in terms of glance behaviour (eg, more anticipatory glances at hazards, fewer long glances off-road). The experienced drivers were significantly better than both novice groups. Hazard approach speed was not significantly different between groups but trended towards better (ie, slower) for trained novices.

Before follow-up, half the trained participants completed ACCEL a second time. At follow-up, differences were observed between trained participants who had completed the programme once versus twice. Participants who repeated the programme were significantly better at anticipating several hazards than those who received it only once. Drivers who completed the programme only once showed decayed effects of training and were no longer significantly different to the control drivers. A major limitation of this follow-up analysis is that it included only 26 drivers, subdivided into three groups, meaning the analyses were underpowered.

3.5.7.2 SAFE-T

The SAFE-T programme integrated training in hazard anticipation, hazard mitigation and attentional maintenance from RAPT, ACT and FOCAL. Although drivers could obviously take the three programmes independently, the benefit of combining them is efficiency: SAFE-T takes only 40–45 minutes total, making it a similar length to each of its individual parent programmes.

Yamani et al (2016) conducted a simulator evaluation of SAFE-T with 48 young novice drivers aged 16–18, who were randomly allocated to SAFE-T, RAPT, or control. Both trained groups made significantly more anticipatory glances towards hazard locations than controls (72–73% vs 47%). SAFE-T drivers also showed superior hazard mitigation skills, as they began moving laterally to avoid a hazard significantly earlier than controls. (RAPT-trained participants moved non-significantly earlier than controls, but the RAPT programme does not actually target this skill.) Finally, SAFE-T participants also showed superior attentional maintenance, making fewer long (≥ 2 seconds) glances off road than both RAPT and control participants. Although this study is limited by its small-scale and short-term evaluation, it provides a useful demonstration that a relatively brief training programme can be as effective as an existing programme such as RAPT, which has previously shown long-term training benefits.

In a further evaluation of SAFE-T, Zhang et al (2018) used personality measures and self-reported driving behaviour data to classify participants as ‘careful’ or ‘careless’ drivers. The evaluation included 48 young drivers aged 18–24, with just over 3 years' driving experience on average, who were randomly allocated to training or control. Analyses of the post-training simulator drive revealed a significant interaction between training condition and driving style: training improved hazard anticipation and attentional maintenance for careful drivers, but not careless drivers. The authors suggest tailoring training programmes to better appeal to different personality traits and driving styles. That said, it is unclear why careless drivers did not improve post-training and therefore what type of changes could improve training effectiveness for careless drivers. Specifically, additional data are needed to assess whether 'careless' driving style impairs the *acquisition* of hazard perception skills during training, or if careless drivers acquire the skills but simply opt not to *apply* them, either because they do not perceive situations as risky or because they actively enjoy experiencing risk.

3.5.8 Commentary training

Commentary training involves having a participant either watch an expert-generated commentary of a driving scene (ie, identifying potential hazards) or self-generate a commentary while driving or viewing a road scene. Several training programmes incorporate commentary training as one component of a larger driver training programme; the studies reported in this section evaluate commentary training as the sole or primary intervention. Most of this research originated in the UK.

Crundall et al (2010) evaluated commentary training with a sample of 40 learner or recently licensed young drivers, who completed simulator drives before training and 2 weeks post-training. Training was conducted on-road, and participants received feedback on their commentary. Compared with controls, the trained group showed a significantly larger reduction in crashes, and greater speed reduction, greater brake pressure and reduced accelerator pressure on approach to hazards. Analysis of hazard subtype revealed the effects were strongest for behavioural prediction of hazards, rather than environmental prediction or dividing and focusing attention. The behavioural hazards involved identifying the presence of a potential hazard and predicting how it might develop, whereas the environmental hazards involved identifying environments that are likely to contain concealed or latent hazards. The divided attention situations involved the presence of multiple potential hazards that must be monitored and assessed to identify the 'true' hazard. There were no significant differences in overall driving speeds, indicating that the training specifically targeted speeds while negotiating hazards. Again, this is an example of training successfully addressing the behaviours targeted in the programme, without necessarily impacting safe driving behaviour more broadly.

McKenna et al (2006) conducted two experiments evaluating training using expert-generated commentary, with samples of 91 and 145 young novice drivers (< 4 years' experience, average age 19 years), randomly allocated to training or control within each experiment. They were tested using video-based driving-related tasks, which required them to make various safety judgements (eg, how fast they would drive, whether it is safe to turn, whether the situation is dangerous). In Experiment 1, trained participants made significantly faster responses in the hazard perception test and indicated that they would adopt more cautious behaviour in the speed and gap acceptance scenarios. Experiment 2 tested whether the lower self-reported speeds in Experiment 1 reflected specific responses to hazards, or a generalised slower/safer approach to driving, by developing a new set of speed test stimuli that included some non-hazardous situations. Trained drivers again reported that they would select significantly slower speeds overall, but the training effect was smaller for less hazardous scenes.

Because different training programmes implement commentary training slightly differently, Wetton et al (2013) directly compared different approaches to commentary training. Young novice drivers (N = 233; mean age 18 years) were randomly assigned to either a control group or one of four training conditions:

1. expert commentary training only
2. expert plus self-generated commentary training
3. ‘What happens next?’ training, in which they watched a clip that was blanked before a hazard materialised, and had 5 seconds to verbalise as many potential outcomes as possible
4. full training, combining expert commentary, self-generated commentary and what happens next.

Participants were tested before, immediately after, and an average of 10 days after training, using a computer-based hazard perception test. Immediately after training, all training conditions showed significantly larger improvement in hazard perception scores than controls, but the effects were largest for full training and smallest for ‘What happens next?’ All conditions showed significant decay at 10-day follow-up, with the ‘What happens next?’ training group no longer being significantly better than controls. This suggests that concurrent commentaries are more effective than anticipation exercises in improving hazard perception. Further evaluation would be needed to confirm that the benefits of commentary training lasted longer than 10 days, and also that they translated into real-world safer driving outcomes.

3.5.9 Other hazard anticipation training programmes

As mentioned, a large proportion of the research on hazard perception training focuses on RAPT, FOCAL, ACT and related programmes in the USA. Several other programmes have been developed in other countries; most focused on hazard anticipation, identifying latent hazards and glance behaviour.

In the UK, which does not have a GDLS, Chapman et al (2002) evaluated a five-stage training programme designed to improve visual search while driving. Participants watched videos of hazardous driving situations and completed progressive tasks, designed to have them search the scene, verbalise what they saw, and try to predict future events. The training took approximately 50 minutes. The evaluation included 143 newly qualified drivers, randomly allocated to training or control, although not all participants completed all phases of the research. Participants completed assessed on-road drives immediately after training and 3–6 months post-training, with visual search, speed and headway measured. Only visual search of dangerous situations showed a significant effect of training; the intervention group showed wider horizontal glance distribution (ie, more scanning left and right).

In Germany, driver training and licensing is more tightly regulated than countries where most hazard perception research has been conducted. Petzoldt et al (2013) therefore developed a computer-based hazard perception training programme for German learner drivers who had completed at least 10 hours of theory and 5 practical lessons. Participants watched simulated videos of traffic scenes, in which hazards could appear horizontally or vertically distant from the driver. The video paused at critical situations and posed multiple-choice questions (eg, ‘Is there a need for action?’). Responses were used to adapt the programme to the driver’s skill level. A paper-based version of the programme was also created, which used static images instead of videos and obviously did not allow for adaptive presentation based on responses to earlier questions. The evaluation included 12 computer-trained, 13 paper-trained and 11 control participants with complete datasets. Participants were evaluated in a simulator drive 2 days post-training. There was a significant effect of training group, with computer-trained participants making critical glance sequences significantly earlier and fixating hazard precursors sooner than controls or paper-trained drivers.

In Israel, Meir et al (2014) developed the video-based Act and Anticipate Hazard Perception Training (AAHPT) to expose young drivers to hazards they would not normally encounter and teach them to anticipate these hazards. The training was targeted for novice drivers (< 3 months’ experience) and included a variety of situations, emphasising scenarios in which hazards did materialise. Three training conditions were included: active, instructional and hybrid. Participants in the active condition had to monitor the scene and press a button when they detected a hazard, whereas participants in the instructional condition received explanations of types of hazards depicted in the videos. The hybrid condition received both elements. The evaluation included 10 young novice drivers (aged 17–18) for each training group, and 10 novices and 21 experienced drivers (aged 23–29) as comparison groups. Training effects were evaluated with a hazard perception test 1-week post-training. Participants in the active and hybrid groups showed greater hazard awareness and greater horizontal spread of eye movements than controls. The instructional condition did not seem as beneficial, with a lot of performance variability within the group and participants often classifying non-hazardous events as hazardous. The authors concluded that the active, practical component is important for improving hazard perception.

In another Israeli study, Kahana-Levy et al (2019) evaluated a shorter hazard anticipation programme similar to the active AAHPT condition. Participants included 21 experienced drivers (~5 years' experience) and 32 novice drivers (~6.7 months experience), with novice drivers randomly allocated to training or control. All experienced drivers received training. Both novice and experienced drivers showed post-training improvements in hazard sensitivity, hazard response and detection of hazard precursors. Unexpectedly, trained novice drivers were significantly better than experienced drivers on several measures. Overall, the results are mixed, which could be because the subgroups were so small. It should also be noted that the 'experienced' group were still relatively young and inexperienced.

In Japan, Shimazaki et al (2017) evaluated HazardTouch, a tablet-based video hazard perception training programme. Traffic crash videos were displayed and paused partway through, so trainees could indicate the location of the hazard. They then watched the remainder of the video to see if they were correct. This intervention only took 10 minutes. The evaluation included 20 drivers, who were randomly allocated to training or control, although the groups were not well matched (significant differences in age, experience and driving frequency). Participants' eye movements were evaluated while they watched a video of a driving scene pre- and post-training. Performance on the pre-training test did not significantly differ between groups. After training, the trained group had a faster time to first fixation on potential hazards and blind spots, and made more fixations on these areas, suggesting superior hazard awareness. As with similar evaluation studies, the conclusions here are limited by the small sample, differences between groups, and lack of long-term evaluation or real-world measures.

In China, Wang et al (2010) compared simulator- and video-based error training for dealing with difficult driving situations. The evaluation involved 32 male novice drivers randomly allocated to one of the two training conditions. Although they only had on average 1 year driving experience, the average age was 23 because Chinese drivers often obtain their licence later than in other countries. The simulator training was experiential: participants were placed in risky situations, in which errors would result in crashes or tickets. The video training was observational, watching videos of the same eight scenarios, with alternative versions depicting correct/safe actions versus erroneous actions resulting in crashes or tickets. After training, participants were scored on errors made during a simulator drive (which was not the same simulator used for training). The simulator-trained group performed significantly better than the video-trained group; both methods yielded near-transfer training benefits, but the simulator group showed superior far transfer. The simulator group also reported lower subjective workload (effort and frustration during the task) and higher intrinsic motivation and metacognition during training. These results are encouraging, but further research would be useful to compare both groups with a no-training control condition (ie, to permit a cost–benefit comparison of the two methods, as the simulator training appears more effective but is presumably more costly). It would also be necessary to evaluate the training in a larger sample, including female drivers.

3.5.10 Hazard perception training for specific contexts

Most hazard perception training programmes aim to improve hazard perception overall or focus on a broad category of hazards such as intersections or latent hazards. However, some programmes target highly specific aspects of hazard perception, such as hazards involving pedestrians, motorcycles, or rural roads.

3.5.10.1 Rural driving

Kumfer et al (2017) designed a ‘supplementary’ computer-based training programme for Texas teen drivers to improve their understanding of the hazards of rural roads. The programme was modelled after existing general hazard perception programmes like Driver ZED and RAPT-3 and was delivered freely online, although all participants completed it at school on a classroom computer. The training programme highlighted key differences between rural and urban driving conditions, including highway driving, curve driving, the impact of adverse weather and other hazards such as animals and work zones. The researchers evaluated the online training programme using a sample of 106 high school students who had completed both pre- and post-intervention surveys, which assessed their risk awareness and typical behaviour both in general driving and specific to rural roads. Students who completed the online programme showed significant increases in knowledge and risk awareness, especially relating to rural roads. The authors noted that prior to the intervention, students had low knowledge about the dangers of rural roads (eg, many were shocked to learn that rural roads are statistically more dangerous than urban roads). This study has obvious limitations in that the sample was much smaller than intended (800 students were initially invited) and self-selected to take training, so it is best viewed as a proof of concept for adapting RAPT-style programmes to specific contexts such as rural driving.

3.5.10.2 Pedestrian hazards

Abele et al (2019) evaluated a training programme specifically focused on pedestrian hazards, which comprised three stages: doing an initial simulator drive, then watching an expert-commentated drive, and finally reviewing their own drive with speed information displayed. The evaluation included 60 male drivers aged 18–24, randomly assigned to training or control. Training did not significantly influence lateral position or self-assessed hazard perception skills, and had only small effects on speed, eye movements and self-assessed driving skills. The effects were also inconsistent, with only 2 out of 7 test scenarios – slower approach and more fixations on hazard – showing positive training effects. The evaluation had some methodological limitations (all testing was conducted within a single session, and the trained group completed more activities than controls), but these would have been more likely to inflate the effect of training. Other similarly brief training programmes (~30 minutes) have yielded positive effects, so it is unclear why this programme did not.

3.5.10.3 Motorcycle hazards

Some research has focused on the problem of car drivers (not) detecting motorcycles. The most common cause of multiple-vehicle crashes involving motorcycles is right-of-way violations, and often in these crashes the car driver responsible claims they ‘looked but failed to see’ the motorcycle. Studies exploring why this occurs are out of scope of the current evaluation, but some studies have evaluated the efficacy of perceptual training to make car drivers more aware of motorcycles, with conflicting results.

Crundall et al (2017) used perceptual training, in which 30 car drivers completed a 6-minute matching game, for which they were randomly assigned to match either motorcycles (training) or fruit (comparison group). This was designed to make drivers more perceptually familiar with motorcycles, so they would be easier to detect in road scenes. After training, participants completed a test that involved detecting cars and motorcycles at intersections, which they had previously completed prior to training. The trained group significantly improved in their detection of motorcycles from pre-test to post-test but did not improve for car detection. In contrast, the control group improved their detection of cars but not motorcycles. A major limitation of the analyses is that they compared the groups individually pre- and post-test but either did not test, or did not report, the direct comparison between training and control groups post-training.

Keyes et al (2019) designed a new motorcycle training programme to address the limitations they noted in Crundall et al's (2017) research. The training programme required participants to search for motorcycles in aerial photographs of car parks. A similar test task was used, in which participants had to detect cars or motorcycles approaching intersections. Forty-two experienced drivers were pseudo-randomly assigned to training or control, but no significant effects of training were detected.

There are several differences between the two training tasks used, so it is difficult to definitively say why one study concluded that training improves perception of motorcycles and the other did not. Keyes et al (2019) hypothesise that their training programme was too brief (< 10 minutes) and would need thousands of trials to demonstrate an effect. Another point to note is that they trained drivers to recognise aerial views of motorcycles, but tested them on frontal views, whereas Crundall et al (2017) used frontal views for both training and test.

A further issue with both tasks is that they assess basic perception, rather than hazard perception and response. Faster and more accurate detection of vehicles could theoretically improve on-road safety by allowing drivers a longer response time, but the drivers must also perceive the vehicle as a hazard worthy of a response (which is better addressed in other training programmes).

3.5.11 Online hazard perception training

Although most hazard perception training is computer-based, often using widely available software such as Microsoft PowerPoint or Flash, surprisingly few studies have evaluated online training programmes.

Isler and Starkey (2012) evaluated eDrive, an online hazard perception training programme with modules targeting visual search, hazard anticipation, responding to risky situations, appropriate speed choice, and commentary training (including both self-generated and expert commentaries). The evaluation compared 634 drivers before and after training, but no sample characteristics were reported. Following training, participants detected significantly more hazards in the online video simulation test, with a moderate increase in hazard detection. They were also significantly faster at detecting hazards, although this effect size was only small in magnitude. The obvious limitation of this research is that it was conducted entirely online, with no control group and no information about the participants or their real-world driving habits.

Arslanyilmaz and Sullins (2019) evaluated game-based multi-user online simulated training (GMOST), an online multi-player game that had participants drive a simulated vehicle through six 'missions'. Players gained or lost points based on their driving performance. The game also incorporated expert commentary videos for hazard situations, which the player had to watch if they did not respond appropriately to the corresponding in-game hazard. Twenty-two teen drivers from local high schools completed the game and then were evaluated in a driving simulator. After training, drivers showed a significant improvement in hazard reaction time (faster) and a non-significant increase in horizontal scanning. Although this study has several significant limitations (eg, small sample, short-term follow-up, no control group, no measures of real-world driving) the training programme itself is notable in that the training could easily be delivered to a large population of novice drivers across multiple geographic locations.

Finally, Zhang et al (2016) evaluated 'Distractology 101', which used a 20-minute online training session to reinforce skills trained during a 30-minute in-person session that involved driving simulator exercises designed to improve awareness of latent hazards and the dangers of secondary task engagement. A group of 839 trained drivers were compared with 26,932 untrained drivers; all were novice teenage drivers

(< 3 years' experience) who had insurance policies with Arbella, the company that created the training. Participants' crash and infraction rates were assessed over a 1-year period. Trained drivers were significantly less likely than untrained drivers to experience a crash (9.6% vs 14.9%) or a violation (19.1% vs 25.7%) and incurred fewer infractions on average (0.25 vs 0.36). Although these results are promising, an obvious caveat is that the sample is non-representative because it consisted entirely of drivers who held an insurance policy in their own name. It is also unclear how trained participants were recruited, but presumably they volunteered and may differ systematically from drivers who chose not to volunteer for the training/research programme.

3.6 Other training programmes

Although most studies identified in the review could be categorised alongside similar training programmes, there were some training programmes that did not fit into any of the previously discussed categories. These are described below.

3.6.1 Feedback training for speed management

Molesworth and colleagues (Krasnova et al 2015, 2016; Molesworth 2012; Molloy et al 2019) conducted a series of studies examining the effect of feedback, described as a cognitive training intervention, to improve the speed management behaviour of young novice drivers. All studies were conducted in Sydney, Australia, with young drivers aged 18–25 who were on their provisional licence. Most used a driving simulator, but Molloy et al (2019) used an instrumented vehicle on public roads.

The main feedback intervention was designed to cognitively involve drivers in the task of speed management. Participants completed a 'paper delivery route' drive (either in the simulator or instrumented vehicle), and afterwards received feedback about their speed performance, including specific consequences for the extent of speeding they exhibited (eg, fines, demerit points, potential harm to others). Across multiple studies, this intervention resulted in lower maximum speeds and less time spent speeding, with the effects maintained – or even enhanced – at 1-week follow-up.

Krasnova et al (2016) separately tested different aspects of feedback, comparing different types of feedback (eg, combined, performance only, performance and finance, performance and safety) and found that all types of feedback had a positive impact on speeding.

Molloy et al (2019) extended the research by examining performance on-road, not just in simulated driving, and comparing concurrent versus post-drive feedback. Seventy-five young novice drivers were randomly allocated to either the control group or one of four training groups:

1. 'Auditory alert feedback' participants received a simple auditory alert during the training drive whenever they exceeded the posted speed limit.
2. 'Combined feedback' participants received post-drive feedback about their number of speed exceedances, maximum speed, and the potential penalties for their speeding (demerit points, fines).
3. 'Self-explanation' participants were asked to complete a written self-reflective exercise about their own speeding behaviour (did they speed and why, potential consequences and alternatives).
4. 'Combined feedback plus self-explanation' participants completed both elements.

The results indicated that the two combined feedback conditions showed the greatest improvement in speed management behaviour, whereas participants assigned to do self-explanation alone showed a smaller improvement. In contrast, control and auditory alert participants actually increased their speeding over successive drives.

Taken together, the results of this research programme suggest that personalised feedback is useful for novice drivers (possibly because their self-assessments are inaccurate). However, any feedback must be contextualised, as in a combined feedback intervention (eg, you exceeded the speed limit on *x* occasions during that drive) rather than concurrent feedback at the time of speeding.

3.6.2 Video-based training for procedural skills

Freeman et al (2015) evaluated a video-based training intervention for reducing run-off-road crashes. This is notably different to the video-based hazard perception programmes in that it is addressing a procedural skill by instructing drivers how to avoid and respond in a run-off-road situation. The videos depicted run-off-road situations in three contexts (highway, curve, rural residential road) and provided detailed instructions on how to recover from a potential run-off-road situation, including recommended speeds. To evaluate the training programme, 75 participants completed two simulator drives (pre- and post-training) that included the same three potential run-off-road situations. Participants were aged 18–36 with an average of over 5 years' driving experience and were randomly allocated to either the training group or the control group. Participants in the training group showed significant improvement from pre- to post-test, with significantly fewer spinouts on both the highway (70% reduced to 16%) and curve driving (50% reduced to 30%) scenarios. The control group did not show any significant improvement.

Although the results appear encouraging, and the idea of video-based training for procedural skills training is interesting, there are some significant limitations and caveats (in addition to the standard limitations that apply to a relatively small sample experimental evaluation). First, all evaluation components were conducted within a single testing session, which likely inflated the effects of training, especially given the fact that training and test scenarios were closely matched. Second, it is unclear whether run-off-road events in the simulator are experienced in the same way as real run-off-road events (ie, the physical and emotional sensations are likely to be different on a real road).

3.6.3 Resilience training

Senserrick et al (2009) evaluated a 1-day workshop focused on building resilience, compared with a more traditional 1-day workshop focused on driving risks. Although the programmes target attitudes rather than skills, this study was included because the evaluation examined official crash and violation rates.

Participants were 2,216 individuals from the DRIVE cohort study in New South Wales, Australia, which tracked over 20,000 young drivers. The resilience programme focused on improving driver safety indirectly by equipping young people with the skills to avoid risky situations such as potential drink driving incidents (eg, by resisting peer pressure). Crash and violation rates were monitored for 2 years; neither workshop impacted violation rates, and the traditional driver-focused workshop did not influence crash rates. The resilience-focused workshop was associated with a 35% reduction in raw crash rates, which increased to a 44% reduction after adjusting for confounding factors. A major limitation of this study is that the two workshops were conducted in different regions, which differed dramatically in socio-economic status. The resilience programme was delivered in lower-income rural areas, whereas the driver-focused programme was delivered in higher-income urban areas. Further research would be required to confirm the safety potential of resilience training.

3.6.4 Team-based training for peer passengers

Because peer passengers have been consistently identified as a key risk factor for young drivers, Lenné et al (2011) trialled a team-based training intervention to improve passenger safety. Thirty-one friend pairs were recruited; all were males aged 18–21 who held a probationary (restricted) driver licence. The research was conducted in Victoria, Australia, where drivers must be at least 18 years old to obtain their probationary licence. Pairs were randomly allocated to receive either team-based training or control, and within pairs the participants were randomly allocated to be either driver or passenger in the post-training simulator evaluation drive. The aim of the training programme was to foster teamwork and communication between peers. It was structured as a 2-hour discussion-based workshop. Discussion points included passenger-related risks for young drivers, responsibilities of drivers and passengers, team concepts (eg, shared goals) and how to apply these to their passenger–driver relationship, and effective communication strategies (ie, to be positive influences).

After training, the pair completed a simulator drive as a driver–passenger dyad. Trained participants showed significantly better performance on several outcome variables (but not all variables). Trained pairs spent less time talking, and trained passengers made fewer unsafe comments than controls. Trained drivers showed significantly longer headway in a vehicle-following task, and earlier braking on approach to an intersection that had an emergency vehicle needing to bypass. These safer behaviours appeared to be genuinely attributable to training, because trained participants self-reported *more* undesirable behaviours (violations, lapses and mistakes) in their everyday driving, as measured by the Driver Behaviour Questionnaire. These results suggest that for at least some drivers, passengers can be trained to facilitate safer driving behaviour. However, it is unclear how long these training effects would last, and whether they would work for all drivers or only a subset (eg, evaluation participants likely had some interest in safe driving, because they volunteered for a driver training study).

3.7 Summary of results

The systematic review found a large number of studies evaluating driver training interventions, of which nearly half focused on hazard perception training. The other half included a diverse range of training methods, with enormous variety in the length of training, length of follow-up, outcome measures reported and sample size studied.

To facilitate comparison of broadly similar studies, Tables 3.1, 3.2 and 3.3 present summaries of evaluations relating to government-endorsed training programmes that are explicitly linked to the licensing process, on-road training programmes and simulator training programmes, respectively. The tables for different training types include slightly different information, as some information is more relevant for specific types of studies (eg, for on-road programmes it is more relevant to include sample size because this varies greatly). There is no equivalent table for hazard perception studies, because the sheer number of studies in this category would make the table too large to be useful, and because similar studies have already been grouped and discussed together in section 3.5.

Table 3.1 Summary of evaluations for driver training programmes incorporated into the licensing process (both pre- and post-licence)

Evaluation	Location	Type of training	Stage & role in licensing process	Key findings
Katila et al 2004	Finland	Skid training	Compulsory; pre-licence	• No significant change in slippery crash rates
Cartensen 2002	Denmark	Linked theory and practical training, including defensive driving and hazard perception	Compulsory, pre-licence	• Significant reduction in multi-vehicle and manoeuvring crashes in first 18 months of licensure • No significant differences after 18 months
Potvin et al 1988	Quebec, Canada	Approved course: • 30 h classroom • 8–10 h in-vehicle	Compulsory for all ages, pre-licence (previously compulsory for 16–17-year-olds only)	• No significant change in licensing or crash rates for males • Significant increase in early licensure (at 16–17 years old) for females • Significant increase in crashes for females aged 16–17 • Significant increase in non-injury crashes for females aged 18+ in the period 6–12 months post-licence
Hirsch et al 2006	Quebec, Canada	12 h practical lessons	Time discount at learner stage; reduced from 12 to 8 months	• No significant difference in overall violation or crash rates in first 450 days of licensure • Significant increase in crashes and violations for drivers who were motivated by time discount, but not for drivers motivated by learning
Boase and Tasca 1998 Zhao et al 2006	Ontario, Canada	Approved course: • 25 h classroom • 10 h in-vehicle	Time discount at G1 (learner) stage; reduced from 12 to 8 months	• Significant higher unadjusted crash rates for trained drivers • No significant differences in self-reported crashes during G2 (restricted) phase after accounting for exposure and experience
Begg and Brookland 2015	New Zealand	Approved defensive driving course	Time discount at restricted stage; reduced from 18 to 12 months	• Drivers who accepted the time discount had significantly higher risk of a traffic offence after progressing to full licence
de Craen and Vlakveld 2013	Netherlands	Intensive (~2 weeks) vs distributed (several months) training	Pre-licence alternative training	• Intensive group significantly more likely to report ‘incidents’ during first 2 years of licensure

Page et al 2004 Freydier et al 2016	France	Supervised driving apprenticeship (3,000 km)	Pre-licence alternative training	• No significant difference in crash likelihood in first 2 years of licensure • Simulator evaluation showed superior speed regulation and lateral position for trained drivers
Shell et al 2015	Nebraska, USA	Certified course: • 20 h classroom • 5 h in-vehicle	Learner stage; removes need to log 50 h of supervised practice	• Significantly fewer crashes and violations in trained group during first 2 years of licensure
Mayhew et al 2014	Oregon, USA	Approved course: • 30 h classroom • 6 h on-road driving • 6 h in-car observation	Pre-licence; reduces supervised practice requirement from 100 h to 50 h	• Significantly higher unadjusted crash rates for trained drivers • Significant reduction in crashes for trained drivers in first 18 months of licensure (but not after that), even after accounting for length of licensure • Significant reduction in violations for trained drivers for first 54 months of licensure
Mayhew et al 2014	Manitoba, Canada	Approved course: • 30 h classroom • 8 h driving • 8 h in-car observation • Plus 24 h supervised practice at home	Pre-licence; can obtain learner's permit 6 months early (at 15.5 years old instead of 16 years old)	• Trained drivers self-reported higher skills, lower exposure • Simulator evaluation showed fewer errors for trained drivers on their learner's permit • Licensing exam showed no significant difference in pass rates, but significantly higher scores for trained drivers

Table 3.2 Summary of evaluations for driver training programmes incorporating on-road training

Evaluation	Location	Training duration	Sample size	Sample characteristics	Key findings
Lund et al 1986 Stock et al 1983	Georgia, USA	SPC: 70 h PDL: 21 h	16,338	High school student learner drivers, randomly allocated to group (SPC, PDL, control)	• Training accelerated licensure; after adjusting for this: – SPC: significantly higher crashes and violations than controls – PDL: no significant differences in crashes or violations
Farmer and Wells 2015	Maryland, USA	90 min	5,279	Learner drivers (aged 16–17) who had completed basic driver education	• No clear effect on crash or violation rates (some analyses showed increases, others showed decreases for the same measure – depending on how the control group was defined)
Gregersen 1996	Sweden	30 min	58	Learner drivers, randomly allocated to traditional or insight skid training	• No significant difference in objective skill level between groups • Insight group significantly less likely to overestimate skid handling skills
Gregersen 1994	Sweden	12 × 45 min	1,793	Learner drivers, pseudo-randomly allocated to group (experimental or traditional training)	• Trained drivers in the experimental programme showed a significant increase in self-reported crashes in the first year of licensure, but a significant decrease in the second year
de Craen et al 2005	Netherlands	1 day	127	Young novice drivers, randomly allocated to group (training, control)	• Inconsistent results – training associated with positive effects at one site, negative at the other – due to practical issues that interfered with evaluation
Molina et al 2007	Spain	1 day	238	Young novice drivers, randomly allocated to group (training, control)	• Significantly higher self-reported careful driving among trained drivers at 9-month follow-up • No other significant differences
Stanley and Mueller 2010	Montana, USA	1 day	347	Young novice drivers who had previously completed basic driver education, randomly allocated to group (training, control)	• No significant differences in official crash or violation rates • Inconsistent differences in self-reported data; trained drivers less likely to report violations in years 1–2, but more likely to report single-vehicle crashes in year 3
Brijs et al 2014	Belgium	3.5 h + 3 weekly emails	366	Young novice drivers, randomly allocated to group (training, control)	• Large significant effect of training on risk knowledge, maintained at 2-month follow-up • Some small significant positive effects of training on attitudes, norms and intentions, but not maintained at follow-up

Evaluation	Location	Training duration	Sample size	Sample characteristics	Key findings
Usami et al 2016	Italy	7 h	15,041	Licensed drivers of varying experience, self-selected to training condition	• Significant increase in violation rates over time for training group (decrease for controls) after accounting for pre-training levels of violations
Petersen et al 2006 Petersen et al 2008 Petersen and Barrett 2009	Queensland, Australia	2-day	39	Licensed drivers of varying experience, method of group allocation not stated	• Training improved postural stability, cornering control, control when changing lanes, smooth braking • Training increased stopping time in vehicle with an anti-lock braking system
Stanton et al 2007	England	8 weekly sessions	75	Experienced drivers	• Personalised coaching improved knowledge, skills and attitudes relating to hazard identification and handling • Significant improvements in vehicle handling and situation awareness for training but not control groups • Attitudes shifted slightly, more internal (vs external) locus of control
Isler et al 2011	New Zealand	1-week	36	Novice drivers aged 15–18	• Vehicle skills training significantly improved vehicle handling but not hazard perception, attitudes or confidence • Higher-order skills training significantly improved hazard perception, visual search and attitudes, and decreased driving confidence

Table 3.3 Summary of simulator training programme evaluations

Evaluation	Type of training	Type of simulator	Sample size	Sample characteristics	Key findings
Cox et al 2009	Experimental, standardised	High fidelity	20	Learner teens (0 hours driving)	• Simulator-trained drivers significantly better than controls in on-road driving assessment (short-term)
Allen et al 2011	Experimental, standardised	Low fidelity	67	Learners – high school driver education students	• Simulator-trained drivers showed significantly better performance than controls on most measures of simulated driving (short-term)
Allen et al 2012	Experimental, standardised	Three variants; low to high fidelity	554	Learners recruited through high schools and licensing offices	• Drivers trained on simulator with wide field of view (135°) had significantly lower crash rates in first year of licensed driving than controls or learners trained with narrow field of view simulator
Rosenbloom and Eldror 2014	Real-world programme	Medium/high fidelity	280	Learners – driving school students	• No clear advantage or disadvantage to simulator training
Hirsch and Bellavance 2017	Real-world programme	High fidelity	1,120 trained	Learners – trainees at driving schools; comparison group from broader population (> 180,000)	• No significant difference in crash rates during first 2 years of licensure • Simulator-trained drivers had fewer infractions than comparison group; only significant for females in period 18–24 months post-licence, and males 6–24 months
Ekeh et al 2013	Experimental, standardised	Medium fidelity	35	Recently licensed teenagers	• No significant differences in crash or violation rates in first 18 months post-licensure
Zhao et al 2019	Experimental, standardised	High fidelity	44	Experienced licensed drivers (10 years' driving)	• Simulator training significantly increased safe driving habits; effects reduced but still significant 3-days post-training
van Leeuwen et al 2013	Experimental, standardised	High fidelity	30	Licensed drivers with just over 6 months experience	• No effect of training
Biondi et al 2020	Experimental, personalised	Medium/high fidelity	69	Licensed drivers with at least 1 year experience	• Precision feedback training significantly improved optimal lane positioning and reduced standard deviation of lateral position • No difference between auditory-only and redundant auditory-visual feedback training
Ka et al 2020	Experimental, personalised	Medium/high fidelity	21	Novice drivers with less than 2 years' experience	• Personalised training significantly improved nearly all safety measures, but no comparison group was included
Lang et al 2018	Experimental personalised vs standard	Virtual reality	50	Licensed drivers with 1–20 years' experience	• Personalised training group showed the most 'good habits' and fastest emergency braking response time, maintained 1-week later • Standardised training showed benefits immediately post-training but not at follow-up

4 Discussion

The current project provided an up-to-date systematic review of the driver training literature. Although many new studies have been published in the past decade, most of these have involved experimental evaluation of hazard perception training or simulator training. This means that the significant limitations and problems with the driver training literature, raised in previous reviews (eg, Beanland et al 2013), have not yet been addressed, and it is still difficult to draw definitive conclusions about the effectiveness of driver training.

4.1 Training as part of the licensing process

Many jurisdictions outside New Zealand include either compulsory or optional training as part of their licensing process, with programmes either created or explicitly endorsed by government bodies. These programmes have yielded inconsistent results, including a mix of negative outcomes, positive outcomes, and programmes with no clear effect on safety outcomes.

It is problematic to compare programmes between jurisdictions because of vast differences in the structure of different licensing systems. For example, some countries only permit learners to undertake supervised practice with a qualified driving instructor, whereas others allow supervision by most fully licensed drivers. Some jurisdictions mandate a minimum number of hours of supervised practice during the learner phase, but the specific number varies greatly (and there is debate over the appropriateness of various thresholds). The current New Zealand GDLS is probably most comparable to GDLS systems used in various USA states and Canadian provinces, some of which also offer (or have previously offered) some form of time discount following completion of an approved training or education course.

In cases where training programmes have been associated with negative outcomes, such as increased crash or violation rates for trained drivers, this is almost always because training was correlated with earlier licensure. Some programmes and policies explicitly facilitate this, through formal time discounts. Other programmes have indirectly encouraged earlier licensure. For example, in Quebec changing the law to mandate driver training for licence applicants of all ages, instead of only requiring it for 16–17-year-olds, inadvertently encouraged female drivers to get their licence at an earlier stage (Hirsch et al 2006). Some optional programmes have also encouraged earlier licensure, such as the DeKalb study; in this case, it is likely that trained drivers felt ‘ready’ to sit their test following completion of the course and therefore attempted the licence test earlier than they otherwise would have (Lund et al 1986).

Most time discount programmes have been associated with negative outcomes. Some evaluations of time discount policies have explicitly asked participants what motivated them to complete the training course and the overwhelming response is to receive a time discount (Begg and Brookland 2015; Hirsch et al 2006). When comparing drivers who are motivated to receive a time discount or other external benefit, such as a reduced insurance premium, with drivers who are motivated to learn, training is unsurprisingly less effective for drivers with external opportunistic motivations (Hirsch et al 2006). Experimental research has also indicated that the effectiveness of training programmes varies depending on drivers’ personality and driving style, with careful drivers being more responsive to training (Zhang et al 2018). This suggests that new, different approaches are required to train less careful and less safety-motivated drivers.

A fundamental problem with trying to develop training-effective approaches is that the so-called ‘careless’ or opportunistically motivated drivers, who typically have poorer safety outcomes, are less likely to volunteer for training or as research participants. Indeed, some of the evaluations reported in this review offered free training to large numbers of participants – several hundred or even thousands – but struggled to attract participants, with less than 20% of those invited willing to complete the training (de Craen et al 2005; Farmer and Wells 2015). This suggests that most young drivers are not sufficiently motivated to complete additional training, even if there is no cost to them, unless there is an obvious perceived benefit. This lack of motivation could signal a fundamental lack of interest in improving driving skills, but could also reflect other factors such as the inconvenience or opportunity cost of undertaking training (eg, if it would require taking time off paid employment, or travelling to a different town or city).

Among the various driver training programmes that were explicitly linked to local licensing systems, the one that appeared to have the most positive effect on safety outcomes was in Nebraska, USA (Shell et al 2015). This training programme was conducted at the learner stage, and rather than providing a discount, it removed the minimum 50-hour supervised driving requirement. Importantly, Nebraska allows learners to undertake supervised driving with lay supervisors. This means that the training programme did not necessarily *replace* supervised driving (as was the case in other evaluations) but rather for many learners was likely completed *in addition to* supervised driving. Unfortunately, drivers in the training group were not required to log their practice hours, so this is speculative, and other studies have suggested that trained drivers undertake fewer hours of supervised practice (Hirsch et al 2006).

4.2 On-road training

Most on-road training programmes that are not part of the licensing process are aimed at post-licence drivers. A significant limitation of studies evaluating these programmes is that almost all involve some degree of sampling bias. This can manifest in two ways:

1. Where programmes are voluntary and participants self-select to sign up for training, there are typically systematic differences between trained and untrained drivers. Some studies address this using time-lagged designs, in which the control/comparison group comprises a subsample of the research volunteers, who complete all evaluation exercises before training and receive the training programme later, after the ‘post-training’ follow-up.
2. Even where studies use randomised controlled designs, there can be differential dropout between groups. Some studies have higher dropout rates for the training group because the training requires greater time commitment from participants, who are unable (or unmotivated) to complete all the activities required. Conversely, other studies have higher dropout rates for the control group because they are less invested or involved in training, and therefore are less motivated to return for follow-up assessments.

That said, although most post-licence on-road training evaluations probably use biased samples who are more interested in or motivated to undertake driver training, few studies have demonstrated positive safety effects of these programmes. This is arguably the opposite of what we would expect with a biased sample (ie, the training should be *more* effective among drivers who voluntarily opt-in), and therefore raises significant concerns about the quality and effectiveness of training programmes.

A key point worth noting here is that most post-licence training programmes are relatively brief (1–2 days) and do not incorporate an exit test assessing skill improvement. These short training programmes may do more harm than good, because drivers believe they have been ‘up-skilled’ and therefore feel better equipped to handle risky situations in future, when this may not be true (especially after a delay, when the benefits of training have decayed). Therefore, on-road training programmes should include aspects of ‘insight training’ designed to help drivers better understand their limitations and avoid overconfidence (Gregersen 1996).

4.3 Simulator training

Relatively few studies have evaluated simulator training (especially in comparison with the large number evaluating hazard perception training). This may reflect issues with the availability of simulators and the difficulties in recruiting participants to attend training at locations that have simulators.

Notably, driving simulators vary massively in terms of their physical setup, visual realism, and programming capabilities (Wynne et al 2019). Researchers refer to simulator fidelity, meaning realism: low-fidelity simulators usually consist of a single computer monitor depicting the forward roadway and game-like controls, whereas high-fidelity simulators include a full vehicle with a motion base and a large visual display (usually on a wraparound projector screen). Medium-fidelity simulators vary in their construction, but notably include a large visual display that can provide a relatively wide field of view that shows the periphery of the driving environment as well as the central areas.

Many simulator training studies have found no significant effect of training, but some have demonstrated positive effects. Ideally, simulator training should be adaptive, providing personalised feedback and tailoring exercises and scenarios based on the individual driver's weakness. This personalisation is arguably the key advantage to using simulators instead of on-road practice, and has been demonstrated as beneficial in several studies, as highlighted in the current review.

Simulators used in training should also provide a wide field of view (at least 135°). This does not mean that only expensive high-fidelity simulators can be used; it is possible to achieve a wide field of view with a low- to medium-fidelity simulator that uses three monitors to represent the forward roadway.

4.4 Hazard perception training

There is an increasingly large literature on hazard perception training; however, most studies involve short-term evaluations of small samples with outcome measures that are indirectly associated with on-road safety.

Many researchers claim that hazard perception reduces crash risk, but in most cases this is an assumption based on indirect evidence: hazard perception test scores are correlated with crash risk, and hazard perception training increases hazard perception test scores, therefore hazard perception training is assumed to reduce crash risk. As noted in this review, relatively few studies have tested this directly, and the effect of training on crash rates is smaller than many would assume based on the large, positive effects often seen for hazard perception scores and glance behaviour (Thomas et al 2016).

That said, hazard perception training programmes do show some potential. The most compelling evidence is for hazard anticipation programmes that train drivers to scan for different types of hazards (eg, RAPT and similar RAPT-inspired programmes), but this could be because these programmes are more established. Other programmes focus on attentional maintenance, specifically minimising eyes-off-road time, and have also shown encouraging results.

Commentary training has also shown some benefits, although this has typically been used with slightly more experienced drivers, and the task of generating a verbal commentary while driving may be too demanding for relatively inexperienced drivers.

Although several promising hazard perception programmes have been developed overseas, especially in the USA, it is worth remembering that effective programmes should be tailored to address the most problematic issues faced by the target population. In other words, hazard perception training courses used in New Zealand should depict local road conditions, and the high-risk scenarios depicted should be designed to represent common crash types experienced by young novice drivers.

4.5 Limitations

As with all research, the current project has strengths and limitations. A strength is that we used a systematic approach to identify relevant research items, including not just published peer-reviewed papers, but also ‘grey’ literature, including technical reports. This is crucial because many driver training evaluations are only published as reports, rather than as peer-reviewed journal articles. This includes some of the larger evaluations of driver training, including the initial analysis of the DeKalb study (Stock et al 1983) and one of the few studies that directly assessed the impact of hazard perception training on crash rates (Thomas et al 2016).

The main weakness is the nature of the research in this area, as most driver training research has significant limitations. These limitations can relate to the driver training programme and/or the evaluation. These limitations have informed a list of recommendations provided later in the report (see section 5.3).

4.5.1 Programme issues

As already noted, driver training programmes vary considerably in their content and duration, which is one reason why previous research has found mixed results supporting driver training as a safety measure.

Furthermore, evaluation reports may provide vague descriptions of training programmes; for example, referring to ‘skid training’ without explaining how it is structured. This type of vague description is problematic because skid training can have positive or negative effects, depending on how it is taught. Some papers had to be excluded from the review altogether because they did not provide any details at all regarding the type of training; for example, studies that broadly classified drivers as ‘trained’ if there was a record of them completing any form of training course. In other cases, only vague details were provided.

Relatively few studies provide comprehensive information about the content of their training, especially if they are reported in academic journal articles (which usually place strict limits on paper length). Fisher et al’s (2017) technical report of the ACCEL programme probably provides the most detailed information about the programme content. This is helpful because others who are looking to implement the programme (or an adapted version of it) know exactly what was involved.

Aside from lack of details, a further issue with many programmes is that the training does not actually improve the skills targeted because, for example, the programme is poorly designed or it is simply too short. As noted throughout the report, several training programmes have been demonstrated to improve specific skills, such as reducing the number of long off-road glances or reducing speed when approaching hazards. However, not all training programmes include an exit test to ensure that programme participants show increased knowledge and/or skills after receiving the training.

4.5.2 Evaluation issues

Although there are many published articles and reports evaluating driver training programmes, not all evaluations are rigorously conducted. Some of the issues are listed below.

- Some evaluations do not include a control group, meaning that observed effects (eg, changes in crash and/or violation rates) could be attributable to factors other than the training programme, such as changes in weather and/or traffic patterns.
  - For example, when evaluating the impact of skid training, it is important to account for seasonal weather and extreme weather because this could affect base rates of slippery road crashes (Cartensen 2002).

  - The impact of weather changes on evaluation results was also highlighted by de Craen et al (2005), who noted that adverse weather in one of the pre-test sessions effectively inflated scores, making it appear that training degraded driver skill.
- In evaluations that do include control/comparison groups, drivers are often not randomly allocated to the training or control group. This means there could be systematic differences between the groups, which could influence the findings.
  - Some researchers attempt to address this by including confounding factors (eg, age, gender, socio-economic status, duration of licensure) in statistical analyses. This is an imperfect approach because it is unclear which factors need to be accounted for.
  - As shown by Usami et al (2016) and Farmer and Wells (2015), adopting different criteria for selecting or matching control group participants can dramatically affect the conclusions of an evaluation, including changing observed effects from being significantly positive to significantly negative outcomes.
- In evaluations that do include control/comparison groups, as mentioned earlier, there may be differential dropout rates for the training group compared with the control group. This again could cause systematic differences between groups, which could influence the findings.
- Some evaluations provide only short-term evaluation of programme effects. Programmes may yield short-term benefits that quickly 'fade', so it is important to evaluate longer-term effects too. This is also useful because some studies have shown that training effects increase over time, or that training appears to have a delayed effect on behaviour and safety.
- Some evaluations use sample sizes that are too small to demonstrate statistically significant effects, which could lead to researchers prematurely concluding a programme is ineffective when it is beneficial (or harmful!)
  - Small samples are especially problematic if the outcome measure of interest is crashes, because crashes are rare events and therefore very large samples are required to detect statistically significant changes in crash rates.
  - Even if the evaluation uses non-crash outcome measures that have greater variability (eg, hazard perception scores, violations) small samples may still limit the generalisability of results. This is because the sample is likely to be non-representative, and a different sample of drivers could yield quite different patterns of results even if they receive the same training. For example, de Craen et al (2005) found opposing effects when comparing two groups of young drivers trained with the same programme at different driving schools.

Notably, all of these evaluation issues have been highlighted in previous research and some problems have been raised repeatedly for decades. The reason these issues persist is primarily that it is expensive and logistically challenging to conduct well-designed large-scale evaluations of driver training; therefore, most studies recently have involved smaller-scale evaluations of experimental programmes.

5 Conclusions and recommendations

Overall, the major conclusion from this research is that driver training is often poorly designed, poorly implemented and poorly evaluated. This has resulted in many evaluations showing either no clear impact of training on safety outcomes, or significant negative outcomes of training whereby trained drivers exhibit higher rates of crashes and violations.

Considering the review findings, the following recommendations are made.

5.1 GDLS time discount

The restricted licence phase time discount following completion of an approved driver training course should be removed. Several evaluations, including one of the New Zealand programme (Begg and Brookland 2015), have suggested that time discounts can do more harm than good (eg, by increasing violation rates or crash risk).

This policy change is likely to be unpopular among at least some young drivers. If the policy is changed, communications outlining the policy change should clearly explain why the time discount is being removed.

It would also be worth considering any broader societal implications of this change, specifically the fact that drivers will need to be at least 18 years of age to obtain a full licence and whether this will have any impact on employment prospects. Assuming other aspects of the licensing system remain unchanged, employment is the aspect most likely to be impacted, rather than mobility and access more generally, because individuals will still be able to drive unsupervised but will have to wait longer to gain a full licence. However, there may be other impacts (eg, social) relating to the fact that drivers will be subject to peer passenger and night driving restrictions for longer. It is unclear whether these impacts would be positive or negative. If the time discount is removed, it would be advisable to monitor the effects of its removal for an extended period, on both driver safety (eg, crash rates and violations) and broader social issues, including access to employment.

5.2 Minimum licensing requirements

Although the current review did not directly address or evaluate the components of different GDLS programmes, it is notable that most overseas jurisdictions have more rigorous licensing requirements than New Zealand. These include:

- longer duration of the learner stage in most jurisdictions (8–12 months vs 6 months in New Zealand)
- longer duration of the restricted stage equivalent in some jurisdictions (2–4 years vs 12–18 months in New Zealand)
- mandatory structured training for learner drivers
- a requirement to undertake a minimum amount of supervised practice during the learner phase.

Because the current project did not systematically evaluate different GDLS programmes, we cannot make specific recommendations about whether any of these policies should be adopted. Rather, we are noting that substantial differences exist in licensing requirements, and we recommend these be reviewed as part of any policy change intended to improve the safety of young novice drivers.

5.3 Best practice for driver training programmes

Because of the significant limitations noted in relation to both programme quality and evaluation quality, it is difficult to make 'best practice' recommendations for future driver training programmes. There are too many open questions remaining, which would need to be addressed through further research. Unfortunately, conducting high quality research is extremely challenging, and even when research teams design an appropriate study, practical realities can interfere with its implementation. This is highly likely to occur in large-scale evaluations of driver training because researchers will have less control and oversight of important aspects of the evaluation (eg, how the training is actually being implemented by instructors and students, timing of assessment components).

With that major caveat, the following recommendations can be made based on previous findings, firstly relating to the content of driver training programmes.

- Driver training programmes should have a clearly articulated aim; that is, a specific issue or skill set they are intended to address.
- Training programmes should be experiential rather than based solely on observation, instruction or theory. This allows students an opportunity to directly experience critical situations, which could improve their skills. Evaluations directly comparing active, experiential programmes with more passive observation and instruction have shown superior performance for the active programmes.
- Training programmes should be designed so that they highlight drivers' limitations, rather than being framed as 'skill improvement'; for example, making drivers aware of the extent of their visual field, the impact of obstructions and blind spots, or their inability to manoeuvre successfully out of certain critical situations.
- Training should take into account individual differences in personality, driving style and motivations for driving. This may require development of entirely new programmes tailored to drivers who have intrinsically low safety motivation.
- Training programmes should involve distributed practice, rather than intensive (massed) practice, as distributed practice has been shown to be more effective for long-term skill/knowledge retention.
- Training programmes should not enable earlier licensure.
- The content of training programmes should be planned out in detail and compared with previous programmes to ensure that they do not inadvertently include components that have been demonstrated as ineffective or even harmful in previous research.
- The content of driver training programmes should be recorded. This does not mean that programmes need to be standardised, but rather that the process for personalising programmes needs to be described in detail. This will help trainers to deliver programmes consistently and will help future researchers and policymakers decide how to design effective programmes (because they will have a more comprehensive understanding of programmes that have been evaluated in the past).
- Training programmes should include appropriate tests before and after training to determine whether the training programme is effective at improving the targeted skills. This should occur even when the training programmes are not being formally evaluated, as it will help identify cases where training has been unsuccessful.

In addition, existing and new training programmes should be evaluated to ensure their ongoing effectiveness. The following recommendations are made for future evaluations of driver training programmes.

- All training evaluations should include an appropriate comparison group.

- Where compulsory training is introduced, the most appropriate comparison is probably an earlier cohort of drivers. However, other changes are likely to occur between cohorts (eg, different weather, new laws or technology) and therefore it is necessary to also consider overall trends in the broader driving population.
- Where training is optional, the comparison group should involve a group of drivers who have not received the training programme.
  - Ideally participants in an evaluation should be randomly assigned to be in either the training group or the control group.
  - One option is to use a 'time lagged' design, in which all participants eventually receive the training, but half of them receive it later, after the evaluation has concluded. However, this option obviously prohibits conducting long-term follow-ups.
  - To avoid systematic differences between groups (eg, volunteer bias, differential dropout rates) the control group should be required to complete an activity that requires similar commitment/effort as the training programme, but without the same skill development component. This is referred to as an 'active' control group and is especially important for experimental or preliminary evaluations of training programmes.
- Where it is not possible to recruit an active control group – for example, in population evaluations of optional programmes incorporated in the licensing process – it is important to examine the influence of variables that may be confounded with training. These may include, but are not limited to:
  - duration of licensure
  - driving exposure
  - education (and parental education)
  - socio-economic status
  - employment status
  - gender
  - ethnicity
  - any other available data, such as performance in the training course, number of hours of supervised training, and number of attempts at the licensing test.
- The report describing the evaluation should contain details about how participants were recruited or selected, how they were assigned to training versus control groups (if applicable), and the number of participants who dropped out or declined to participate (if applicable).
- The evaluation should consider, as separate analyses, whether the training improved the targeted skills (ie, if it is a hazard anticipation course, did it improve hazard anticipation in a test drive?) and whether the training improved safety.
- Evaluations should include analyses of statistical power. These can be conducted either before or after running the study, and there are freely available computer programs and applications (eg, GPower). A statistical power analysis conducted before beginning the evaluation would provide guidance on the required sample size, based on the anticipated effect size. A posthoc power analysis conducted after the study could help identify whether analyses were underpowered (eg, if the final sample size was smaller than intended).

References

*References marked * are included in the systematic review.*

*Abele, L, S Haustein, LM Martinussen and M Moller (2019) Improving drivers' hazard perception in pedestrian-related situations based on a short simulator-based intervention. *Transportation Research Part F: Traffic Psychology and Behaviour 62*: 1–10.

*Agrawal, R, M Knodler, DL Fisher and S Samuel (2018) Virtual reality headset training: Can it be used to improve young drivers' latent hazard anticipation and mitigation skills. *Transportation Research Record: Journal of the Transportation Research Board 2672*, no.33: pp 20–30.

*Ahmadi, N, A Katrahmani and MR Romoser (2018) Short and long-term transfer of training in a tablet-based teen driver hazard perception training program. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 62*, no.1: 1965–1969.

*Allen, RW, GD Park, ML Cook and D Fiorentino (2012) Simulator training of novice drivers: A longitudinal study. *Advances in Transportation Studies*, Section B 27: 51–68.

*Allen, W, G Park, S Terrace and J Grant (2011) Detecting transfer of training through simulator scenario design: A novice driver training study. In *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Lake Tahoe, CA, June 2011.

*Arslanyilmaz, A and J Sullins (2019) Multi-player online simulated driving game to improve hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour 61*: 188–200.

Beanland, V, N Goode, PM Salmon and MG Lenné (2013) Is there a case for driver training? A review of the efficacy of pre- and post-licence driver training. *Safety Science 51*: 127–137.

*Begg, D and R Brookland (2015) Participation in driver education/training courses during graduated driver licensing, and the effect of a time-discount on subsequent traffic offenses: Findings from the New Zealand Drivers Study. *Journal of Safety Research 55*: 13–20.

*Biondi, FN, R Rossi, M Gastaldi, F Orsini and C Mulatti (2020) Precision teaching to improve drivers' lane maintenance. *Journal of Safety Research 72*: 225–229.

*Boase, P and L Tasca (1998) *Graduated licensing system evaluation: Interim report '98*. Ministry of Transportation of Ontario, Toronto.

*Brijs, K, A Cuenen, T Brijs, RA Ruiter and G Wets (2014) Evaluating the effectiveness of a post-license education program for young novice drivers in Belgium. *Accident Analysis and Prevention 66*: 62–71.

Carstensen, G (1994) Evaluation of a new driver education in Denmark. In *Proceedings of the Conference Strategic Highway Research Program (SHRP) and Traffic Safety on Two Continents*, Hague, Netherlands, September 1993.

*Carstensen, G (2002) The effect on accident risk of a change in driver education in Denmark. *Accident Analysis and Prevention 34*, no.1: 111–121.

*Chapman, P, G Underwood and K Roberts (2002) Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour 5*, no.2: 157–167.

*Cox, CV, R Wharam, R Mouran and DJ Cox (2009) Does virtual reality driving simulation training transfer to on-road driving in novice drivers? A pilot study. *Chronicle of the American Driver and Traffic Safety Education Association 57*, no.1: 9, 21–25.

*Crundall, D, B Andrews, E van Loon and P Chapman (2010) Commentary training improves responsiveness to hazards in a driving simulator. *Accident Analysis and Prevention 42*, no.6: 2117–2124.

*Crundall, D, A Howard and A Young (2017) Perceptual training to increase drivers' ability to spot motorcycles at T-junctions. *Transportation Research Part F: Traffic Psychology and Behaviour 48*: 1–12.

*de Craen, S, J Vissers, M Houtenbos and D Twisk (2005) *Young drivers experience: the results of a second phase training on higher order skills. Evaluation study in the framework of the European project NovEV*. 94 pp.

*de Craen, S and WP Vlakveld (2013) Young drivers who obtained their licence after an intensive driving course report more incidents than drivers with a traditional driver education. *Accident Analysis and Prevention 58*: 64–69.

*Divekar, G, S Samuel, A Pollatsek, FD Thomas, K Korbelak, RD Blomberg and DL Fisher (2016) Effects of a PC-based attention maintenance training program on driver behavior can last up to four months simulator study. *Transportation Research Record 2602*: 121–128.

*Ekeh, AP, K Herman, D Bayham, R Markert, M Pedoto and MC McCarthy (2013) Pilot evaluation of the short-term effect of driving simulation on novice adolescent drivers. *Journal of Trauma and Acute Care Surgery 75*, no.1: 83–87.

*Farmer, CM and JK Wells (2015) *Crash and citation records of young drivers with skid avoidance training*. 23 pp.

*Fisher, DL, R Glaser, NE Laurie, A Pollatsek and JF Brock (1998) Evaluation of PC-based younger driver training program: Use of a driving simulator. *Proceedings of the Human Factors and Ergonomics Society 2*: 1403–1407.

*Fisher, DL, NE Laurie, R Glaser, K Connerney, A Pollatsek, SA Duffy and J Brock (2002) Use of a fixed-base driving simulator to evaluate the effects of experience and PC-based risk awareness training on drivers' decisions. *Human Factors 44*, no.2: 287–302.

*Fisher, DL, AP Pollatsek and A Pradhan (2006) Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention 12*, no.S1: i25–i29.

*Fisher, DL, AK Pradhan, A Pollatsek and MA Knodler Jr (2007) Empirical evaluation of hazard anticipation behaviors in the field and on driving simulator using eye tracker. *Transportation Research Record 2018*: 80–86.

*Fisher, DL, J Young, L Zhang, M Knodler and S Samuel (2017) *Accelerating teen driver learning: Anywhere, anytime training*. 141 pp.

*Freeman, P, DM Neyens, J Wagner, F Switzer, K Alexander and P Pidgeon (2015) A video based run-off-road training program with practice and evaluation in a simulator. *Accident Analysis and Prevention 82*: 1–9.

*Freydier, C, C Berthelon and M Bastien-Toniazzo (2016) Does early training improve driving skills of young novice French drivers? *Accident Analysis and Prevention 96*: 228–236.

*Gregersen, NP (1994) Systematic cooperation between driving schools and parents in driver education, an experiment. *Accident Analysis and Prevention 26*, no.4: 453–461.

*Gregersen, NP (1996) Young drivers' overestimation of their own skill – An experiment on the relation between training strategy and skill. *Accident Analysis and Prevention 28*, no.2: 243–250.

*Hirsch, P and F Bellavance (2017) Transfer of skills learned on a driving simulator to on-road driving behavior. *Transportation Research Record 2660*: 1–6.

*Hirsch, P, U Maag and C Laberge-Nadeau (2006) The role of driver education in the licensing process in Quebec. *Traffic Injury Prevention 7*, no.2: 130–142.

*Isler, RB and N Starkey (2012) Driver education and training as evidence-based road safety interventions. In *Proceedings of the Australasian Road Safety Research, Policing and Education Conference*, Wellington, New Zealand, October 2012.

*Isler, RB, NJ Starkey and P Sheppard (2011) Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. *Accident Analysis and Prevention 43*, no.5: 1818–1827.

*Ka, E, DG Kim, J Hong and C Lee (2020) Implementing surrogate safety measures in driving simulator and evaluating the safety effects of simulator-based training on risky driving behaviors. *Journal of Advanced Transportation 2020*, no.7525721: 1–12.

*Kahana-Levy, N, S Shavitzky-Golkin, A Borowsky and E Vakil (2019) Facilitating hazard awareness skills among drivers regardless of age and experience through repetitive exposure to real-life short movies of hazardous driving situations. *Transportation Research Part F: Traffic Psychology and Behaviour 60*: 353–365.

*Katila, A, E Keskinen, M Hatakka and S Laapotti (2004) Does increased confidence among novice drivers imply a decrease in safety? The effects of skid training on slippery road accidents. *Accident Analysis and Prevention 36*, no.4: 543–550.

*Keyes, H, F Green, C Compton and M Staton (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 Psychology and Behaviour 67*: 57–65.

*Krasnova, O, B Molesworth and A Williamson (2015) The effect of cognitive-based training interventions on driver speed management behavior: A driving simulator study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 59*: 1796–1800.

*Krasnova, O, B Molesworth and A Williamson (2016) Understanding the effect of feedback on young drivers' speeding behavior. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 60*: 1986–1990.

*Krishnan, A, S Samuel, M Romoser and D Fisher (2015) A study to evaluate the effectiveness of a tablet-based training program for younger drivers: Addressing distraction head on. *Proceedings of the Human Factors and Ergonomics Society 59*: 1651–1655.

*Krishnan, A, S Samuel, Y Yamani, MR Romoser and DL Fisher (2019) Effectiveness of a strategic hazard anticipation training intervention in high risk scenarios. *Transportation Research Part F: Traffic Psychology and Behaviour 67*: 43–56.

*Kumfer, W, H Liu, D Wu, D Wei and S Sama (2017) Development of a supplementary driver education tool for teenage drivers on rural roads. *Safety Science 98*: 136–144.

*Lang, Y, L Wei, F Xu, Y Zhao and LF Yu (2018) Synthesizing personalized training programs for improving driving habits via virtual reality. In *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 – Proceedings*, Reutlingen, Germany, March 2018.

*Lenné, MG, CC Liu, PM Salmon, M Holden and S Moss (2011) Minimising risks and distractions for young drivers and their passengers: An evaluation of a novel driver-passenger training program. *Transportation Research Part F: Traffic Psychology and Behaviour 14*, no.6: 447–455.

*Lund, AK, AF Williams and P Zador (1986) High school driver education: Further evaluation of the Dekalb County study. *Accident Analysis and Prevention 18*, no.4: 349–357.

*Madigan, R and R Romano (2020) Does the use of a head mounted display increase the success of risk awareness and perception training (RAPT) for drivers? *Applied Ergonomics 85*, no.103076: 1–9.

*Mayhew, D, K Marcoux, K Wood, H Simpson, W Vanlaar, L Lonero and K Clinton (2014) *Evaluation of beginner driver education programs: Studies in Manitoba and Oregon*. 369 pp.

McCartt, AT, DR Mayhew, KA Braitman, SA Ferguson and HM Simpson (2009) Effects of age and experience on young driver crashes: Review of recent literature. *Traffic Injury Prevention 10*, no.3: 209–219.

*McDonald, CC, V Kandadai, H Loeb, T Seacrist, YC Lee, D Bonfiglio, DL Fisher and FK Winston (2015) Evaluation of a risk awareness perception training program on novice teen driver behavior at left-turn intersections. *Transportation Research Record 2516*: 15–21.

*McKenna, FP, MS Horswill and JL Alexander (2006) Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology: Applied 12*, no.1: 1–10.

*Meir, A, A Borowsky and T Oron-Gilad (2014) Formation and evaluation of Act and Anticipate Hazard Perception Training (AAHPT) intervention for young novice drivers. *Traffic Injury Prevention 15*, no.2: 172–180.

Ministry of Transport (2021) *Safety – Annual statistics. Young drivers*. Accessed 24 February 2021. https://www.transport.govt.nz/statistics-and-insights/safety-annual-statistics/sheet/young-drivers

Moher, D, A Liberati, J Tetzlaff, DG Altman and The PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine 6*, no.7:e1000097.

*Molesworth, B (2012) *Improving drivers' risk management behaviour: an assault on speeding*. 12 pp.

*Molina, J, J Sanmartin, E Keskinen and N Sanders (2007) Post-license education for novice drivers: Evaluation of a training programme implemented in Spain. *Journal of Safety Research 38*, no.3: 357–366.

*Molloy, O, BRC Molesworth and A Williamson (2019) Which cognitive training intervention can improve young drivers' speed management on the road? *Transportation Research Part F: Traffic Psychology and Behaviour 60*: 68–80.

*Muttart, JW, S Dinakar, DL Fisher, TM Garrison and S Samuel (2019) Evaluation of a training intervention to improve novice drivers' hazard mitigation when approaching left turn scenarios. *Transportation Research Record 2673*, no.12: 474–484.

*Page, Y, MC Ouimet and S Cuny (2004) An evaluation of the effectiveness of the supervised driver-training system in France. In *48th Annual Proceedings – Association for the Advancement of Automotive Medicine*, Biscayne, FL, September 2004.

*Petersen, A and R Barrett (2009) Postural stability and vehicle kinematics during an evasive lane change manoeuvre: A driver training study. *Ergonomics 52*, no.5: 560–568.

*Petersen, A, R Barrett and S Morrison (2006) Driver-training and emergency brake performance in cars with antilock braking systems. *Safety Science 44*, no.10: 905–917.

*Petersen, A, R Barrett and S Morrison (2008) Enhanced postural stability following driver training is associated with positive effects in vehicle kinematics during cornering. *Human Factors 50*, no.1: 159–172.

*Petzoldt, T, T Weis, T Franke, JF Krems and M Bannert (2013) Can driver education be improved by computer based training of cognitive skills? *Accident Analysis and Prevention 50*: 1185–1192.

Pollatsek, A, V Narayanaan, A Pradhan and DL Fisher (2006) Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator. *Human Factors 48*, no.3: 447–464. https://doi.org/10.1518%2F001872006778606787

*Potvin, L, F Champagne and C Laberge-Nadeau (1988) Mandatory driver training and road safety: The Quebec experience. *American Journal of Public Health 78*, no.9: 1206–1209.

*Pradhan, AK, G Divekar, K Masserang, M Romoser, T Zafian, RD Blomberg, FD Thomas, I Reagan, M Knodler, A Pollatsek and DL Fisher (2011) The effects of focused attention training on the duration of novice drivers' glances inside the vehicle. *Ergonomics 54*, no.10: 917–931.

*Pradhan, AK, A Pollatsek, M Knodler and DL Fisher (2009) Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics 52*, no.6: 657–673.

*Regan, MA, HA Deery and TJ Triggs (1998) Training for attentional control in novice car drivers: A simulator study. *Proceedings of the Human Factors and Ergonomics Society 2*: 1452–1456.

*Regan, MA, TJ Triggs and ST Godley (2000) Simulator-based evaluation of the DriveSmart novice driver CD-ROM training product. In *Proceedings of the Australasian Road Safety Research, Policing & Education Conference*, Melbourne, Australia, November 2000.

*Rosenbloom, T and E Eldror (2014) Effectiveness evaluation of simulative workshops for newly licensed drivers. *Accident Analysis and Prevention 63*: 30–36.

*Samuel, S, T Zafian, A Borowsky, MR Romoser and DL Fisher (2013) Can young drivers learn to anticipate hidden hazards: a driving simulator study. In *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Bolton Landing, NY, June 2013.

Schiff, R (2019) Evaluation of the Graduated Driver Licensing System. *Transport Evidence Base report 19/1 A*. 91 pp. https://www.transport.govt.nz//assets/Uploads/Report/GDLS-evaluation-report.pdf

*Senserrick, T, R Ivers, S Boufous, HY Chen, R Norton, M Stevenson, E van Beurden and A Zask (2009) Young driver education programs that build resilience have potential to reduce road crashes. *Pediatrics 124*, no.5: 1287–1292.

*Shell, DF, IM Newman, AL Cordova-Cazar and JM Heese (2015) Driver education and teen crashes and traffic violations in the first two years of driving in a graduated licensing system. *Accident Analysis and Prevention 82*: 45–52.

*Shimazaki, K, T Ito, A Fujii and T Ishida (2017) Improving drivers' eye fixation using accident scenes of the HazardTouch driver-training tool. *Transportation Research Part F: Traffic Psychology and Behaviour 51*: 81–87.

*Stanley, LM and J Mueller (2010) *Effects of defensive vehicle handling training on novice driver safety: Phase 3. Data analysis and results*. 49 pp.

*Stanton, N, G Walker, M Young, T Kazi and P Salmon (2007) Changing driver's minds: The evaluation of an advanced driver coaching system. *Ergonomics 50*, no.8: 1209–1234.

*Stock, JR, IK Weaver, HW Ray, TR Brink and MG Sadof (1983) *Evaluation of safe performance secondary school driver education curriculum project. Final report*. 160 pp.

*Thomas, FD, SL Rilea, RD Blomberg, RC Peck and KT Korbelak (2016) *Evaluation of the safety benefits of the risk awareness and perception training program for novice teen drivers*. 68 pp.

*Unverricht, J, Y Yamani, J Chen and WJ Horrey (2020) Minding the gap: Effects of an attention maintenance training program on driver calibration. *Human Factors*.

*Usami, DS, L Persia and V Iurato (2016) Evaluation of Post-license Advanced Driver Training in Italy. *Transportation Research Procedia 14*: 3859–3866.

*van Leeuwen, PM, R Happee and JCF de Winter (2013) Investigating the effect of a visual search task for simulator-based driver training. In *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Bolton Landing, NY, June 2013.

*Wang, YB, W Zhang and G Salvendy (2010) A comparative study of two hazard handling training methods for novice drivers. *Traffic Injury Prevention 11*, no.5: 483–491.

*Wetton, MA, A Hill and MS Horswill (2013) Are what happens next exercises and self-generated commentaries useful additions to hazard perception training for novice drivers? *Accident Analysis and Prevention 54*: 57–66.

Wynne, RA, V Beanland and PM Salmon (2019) Systematic review of driving simulator validation studies. *Safety Science 117*: 138–151

*Yamani, Y, S Samuel, MA Knodler and DL Fisher (2016) Evaluation of the effectiveness of a multi-skill program for training younger drivers on higher cognitive skills. *Applied Ergonomics 52*: 135–141.

*Zafian, TM, S Samuel, J Coppola, EG O'Neill, MRE Romoser and DL Fisher (2016) On-road effectiveness of a tablet-based teen driver training intervention. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 60*: 1926–1930.

*Zhang, T, F Hajiseyedjavadi, Y Wang, S Samuel, X Qu and D Fisher (2018) Training interventions are only effective on careful drivers, not careless drivers. *Transportation Research Part F: Traffic Psychology and Behaviour 58*: 693–707.

*Zhang, T, J Li, H Thai, T Zafian, S Samuel and DL Fisher (2016) Evaluation of the effect of a novice driver training program on citations and crashes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 60*: 1991–1995.

*Zhao, J, RE Mann, M Chipman, E Adlaf, G Stoduto and RG Smart (2006) The impact of driver education on self-reported collisions among young drivers with a graduated license. *Accident Analysis and Prevention 38*, no.1: 35–42.

*Zhao, X, W Xu, J Ma and Y Gao (2019) The PNE driving simulator-based training model founded on the theory of planned behavior. *Cognition, Technology & Work 21*, no.2: 287–300.

Appendix: Description of the New Zealand Advanced Driver Training System

The following information was taken from the Waka Kotahi website[2] to provide additional context for readers unfamiliar with the New Zealand Advanced Driver Training System. It must be stressed, however, that the current report did not set out to evaluate current New Zealand provision.

Advanced driving courses teach you important skills for driving

Advanced driving courses are designed to help you identify and avoid or handle potential hazards you may encounter while driving.

They teach skills that can help reduce the likelihood of you being involved in a crash, such as eye scanning while driving, hazard detection and risk management.

Passing an approved advanced driving course can:

- reduce the minimum age for a full car licence (from 18 years to 17½ years)
- reduce the time you need to stay on a restricted car licence (by either 3 or 6 months depending on your age).

You can do an approved driving course once you have got your learner or restricted car licence.

Waka Kotahi does not administer driving courses. We advise you check the course you are undertaking is relevant to what you want to achieve; eg, if you want to reduce the time you need to hold a restricted car licence, attend one of the courses listed below.

Approved course providers

There are two approved advanced driving courses in New Zealand that can reduce the time you have to hold your restricted car licence before you can get a full car licence. These are:

Defensive driving course

The defensive driving course is provided by trainers who are affiliated to the AA (New Zealand Automobile Association) Driving School.[3]

The course takes 9 hours to complete and includes:

- 4 classroom sessions
- 1 in-car session
- access to the online driver training system.

Street Talk

Street Talk courses are offered through the New Zealand Institute of Driver Educators (NZIDE).[4]

The course includes:

- 4 interactive sessions of 2 hours in classroom
- up to an hour on-road practical session.

On average the courses cover a 2-week period with at least two sessions in each week.

[2] https://www.nzta.govt.nz/driver-licences/getting-a-licence/licences-by-vehicle-type/cars/advanced-driving-course/

[3] https://www.aa.co.nz/drivers/learn-to-drive-the-aa-way/defensive-driving-courses/

[4] https://www.street-talk.co.nz/