Transportation Research Part F: Psychology and Behaviour 94 (2023) 398–412

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

Further down the road: The enduring effect of an online training course on novice drivers' hazard perception skill

Mark S. Horswill [a,*], Andrew Hill [a,b], Lisa Buckley [a,c], Genevieve Kieseker [a], Francine Elrose [a]

[a] School of Psychology, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
[b] Minerals Industry Safety and Health Centre, Sustainable Minerals Institute, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
[c] School of Public Health, The University of Queensland, Herston, Brisbane, QLD 4006, Australia

ARTICLE INFO

Keywords:
Novice drivers
Young drivers
Hazard perception
Situation awareness
Crash risk
Anticipation
Driving

ABSTRACT

Hazard perception training has been proposed as a potential strategy for reducing crash risk in drivers. Although previous work has shown that hazard perception is a trainable skill, it also indicates that the effect of hazard perception training often decays over time, sometimes to the point of being negated. In the present study, we evaluate a hazard perception training course specifically designed to minimize the rate of forgetting through the use of evidence-based strategies such as incorporating distributed practice and taking advantage of the contextual interference effect. We report data on a sample of novice drivers, in which we evaluated the effect of the course on a validated measure of hazard perception skill (a verbal-response "hazard prediction test") in both the short term (M = 9.24 days, or 1.32 weeks) and medium term (M = 9.87 weeks, or 2.27 months), using a randomized control design. At both retention points, trained participants scored significantly and substantially higher than an untrained control group, adjusting for baseline performance. Further, the hazard prediction scores of trained drivers did not significantly decrease between the short- and medium-term assessment points, indicating that there was no evidence of skill decay over this time frame. There was also no evidence that completing the course increased driver confidence at any time point, which is important given that increased confidence could potentially lead to unintended negative effects on crash risk. The results suggest that the training course improves a skill that has been associated with crash risk over a time period that is of practical value, especially as refresher training sessions could be used to further extend the effect into the longer term (i.e., beyond the duration examined in this study) if necessary.

1. Introduction

Hazard perception, which may be defined as a driver's ability to anticipate dangerous traffic situations, has been demonstrated to be one of the few driving skills that can distinguish those who crash from those who do not crash (Horswill, 2016a). Scores on tests of hazard perception skill have been linked with drivers' crash involvement both prospectively and retrospectively (Boufous et al., 2011; Cheng et al., 2011; Darby et al., 2009; Horswill et al., 2010; Horswill et al., 2020; Horswill, Hill et al., 2015; McKenna & Horswill,

* Corresponding author at: School of Psychology, The University of Queensland, St Lucia, QLD 4072, Australia.
E-mail address: m.horswill@psy.uq.edu.au (M.S. Horswill).

https://doi.org/10.1016/j.trf.2023.02.011
Received 26 May 2022; Received in revised form 15 February 2023; Accepted 16 February 2023
Available online 15 March 2023
1369-8478/

1999; Rosenbloom et al., 2011; Shimazaki & Ishida, 2009; Tūskė et al., 2019; Wells et al., 2008). Hazard perception is also a skill that increases with experience (Crundall, 2016; Hill et al., 2019; Sun et al., 2019), and has been linked with key behavioural crash factors, including distraction (Horswill & McKenna, 1999), fatigue (Smith et al., 2009), risk-taking behaviour (McKenna et al., 2006), and impairment (e.g., alcohol; Deery & Love, 1996).

1.1. Hazard perception training

The empirical link between hazard perception and road safety has resulted in calls to develop driver training interventions that target this type of skill (Glendon, 2014; Shinar, 2007), in the expectation that this will lead to reduced crash risk. Consequently, there have been a number of attempts to improve drivers' hazard perception skill via training (Horswill, 2016b; Šeibokaitė et al., 2022; Unverricht et al., 2018). The theoretical approach typically taken by these interventions (either explicitly or implicitly) is to conceptualize hazard perception skill as being linked with drivers' mental representations of the driving environment (Horswill & Hill, 2021). Those who are good at hazard perception are better at perceiving and understanding elements of the driving environment, allowing them to make predictions regarding future states of that environment, including the potential for crashes; that is, they have superior situation awareness (Horswill & McKenna, 2004). Hazard perception training interventions therefore typically focus on improving these mental representations of the traffic environment.

To give an example, Wetton et al. (2013) evaluated the effect of two types of video-based training exercise on novice drivers' hazard perception skill. The first exercise-type ("Commentary Drive") involved trainees watching a video clip of traffic, filmed from the driver's perspective, while commenting on where they would look and what they would monitor to maximize safety. Trainees were then shown the same clip with the addition of a commentary generated by an expert driver. The expert commentary allowed trainees to see how they could improve their own commentary in subsequent exercises. The second exercise-type ("What Happens Next") involved trainees watching a traffic clip that cut to black just before a potentially dangerous situation. Trainees were asked to generate predictions as to what might happen after the cut point. They then heard an expert generating predictions on the same clip, before watching the clip play past the cut point to reveal what actually happened. Participants were randomly assigned to complete either *commentary drive* exercises, *what happens next* exercises, a combination of both exercise-types, or placebo training (i.e. video-based driver training that did not focus on hazard perception). Participants' hazard perception skill was tested before the training, immediately afterwards, and one week later using video-based response-time hazard perception tests. Compared with the placebo control group, trainees' hazard perception skill was significantly improved immediately following the training for both hazard perception training exercises-types used in isolation. However, a significant decay in the training effect was found one week later for each exercise-type to the extent that the training effects were no longer significant. When both exercises-types were used together, the training effect remained significant after a week, but the decay in skill was still significant.

These findings indicate that it is possible to train hazard perception skill but, even when successful, the effects of the training might be short-lived in some circumstances necessitating evaluations that extend to the medium-term.

1.2. Studies examining the durability of hazard perception training

Several studies have examined the effect of hazard perception training in the longer term. When drivers aged 65 and over were trained using a combination of the same two exercise types as Wetton et al. (2013), their hazard perception skill (also measured using video-based response-time hazard perception tests) was improved, and this improvement was maintained for three months without significant decay (Horswill, Falconer et al., 2015). We speculate that this superior retention is a function of trainees' goals, where older drivers might be more motivated to improve their driving ability than novices, given the threat of age-related driving cessation and the substantial negative effects associated with this (e.g., depression, decreased health, and increased mortality; Chihuri et al., 2016).

Chapman et al. (2002) evaluated a video-based training intervention (including *commentary drives*) with novice drivers. The intervention resulted in a greater horizontal spread in drivers' gaze patterns, suggesting more effective hazard scanning, both during real driving and while watching video clips of traffic. In a second assessment conducted three to six months later, this training effect persisted with the video clips but became non-significant for on-road driving. Initial post-training reductions in fixation duration when viewing videos of traffic (also considered to be a marker of effective visual search for hazards) were also not retained in the follow-up assessment.

Taylor et al. (2011) evaluated a computer-based hazard perception training intervention ("RAPT") by measuring novice drivers' fixations on predetermined areas relevant to hazard perception (such as locations that might conceal a road user) during a real drive. Trained drivers were more likely to fixate on such areas than a control group, both immediately and around 8 months later, without significant skill decay. However, for a newer version of the same intervention, Thomas et al. (2017) found that novice driver training effects did decay after one month (measured via the number of hazards detected in video clips), though improvements from baseline remained significant.

Ahmadi et al. (2018) evaluated a tablet-based hazard perception training program also based on RAPT and assessed novice drivers 1 week and 6 months afterwards in an on-road drive. During the drive, trainees' situation awareness was scored based on both a verbal commentary that the drivers produced and their gaze patterns. There was decay found across all measured scenario-types between the two measurement points. However, the training effect remained significant for one scenario-type (crosswalk scanning and decisions) at six months, but not for other scenario-types (turning, 4-way intersections). In contrast to the studies discussed above, Regan et al. (2000) found no evidence of decay when evaluating the effects of another computer-based hazard perception intervention with inexperienced drivers. In that study, drivers who received the training exhibited safer behaviour in a driving simulator on both testing occasions. However, the simulation-based assessments were conducted only three weeks apart (i.e., after one week and four weeks post-training).

More recently, Reyes and O'Neal (2020) reported a study in which they evaluated the effects of two previously established computer-based hazard anticipation training programs (PALM and ACCEL) after six weeks. They found no overall effect of training in a simulated drive (using gaze direction and driving behaviour measures), for either program, despite previous evidence indicating performance improvements immediately after training in one of the programs (ACCEL).

Overall, the evidence indicates that creating hazard perception training that results in retention over a period of months for novice drivers is possible, but is not a foregone conclusion. In some studies, there was no decay in the training effect. In others, there was decay, but the training effect remained significant. In yet others, an initial training effect became non-significant at follow-up. However, differences between the interventions, evaluation measures, samples, and retention durations across the different studies make it hard to draw conclusions.

1.3. Potential strategies for increasing the likelihood of hazard perception training retention

In light of the previous literature on hazard perception training retention, how can we maximize the chances that new hazard perception training interventions will have an enduring effect, without skill decay? One approach is to consider this issue in the context of the forgetting curve, proposed by Ebbinghaus (Baddeley et al., 2020), which highlights the high rate at which individuals forget learned material. There is a considerable literature on strategies for minimizing the forgetting curve (Bahrick & Hall, 1991; Willingham, 2009), which can potentially be used in the driver training context.

One such strategy is to structure learning into a format that incorporates distributed practice (Cepeda et al., 2006), with multiple spaced-out training sessions (Pashler et al., 2007). The breaks between sessions would ideally include the opportunity for sleep, given that long-term learning has been found to be enhanced if trainees sleep between training sessions (Mazza et al., 2016). In this context, it is worth noting that the successfully retained driver training intervention reported by Regan et al. (2000) was spaced over five weekly sessions.

Another strategy associated with improved retention is to take advantage of the contextual interference effect (Broadbent et al., 2017; Magill & Hall, 1990), which refers to the finding that higher levels of interference during learning, usually induced by switching tasks frequently, can lead to improved recall and retention. This suggests that drivers might benefit from having a number of different training exercise types interleaved with one another (Horswill, Hill, Silapurem et al., 2021).

Another possibility is to increase the number of training exercises completed (i.e. the total duration of the training intervention), given that the quantity of deliberate practice is known to be a key factor in successful skill acquisition (Ericsson et al., 1993). We also speculate that trainee motivation is likely to be important to retention, given that the hazard perception skill retention found for

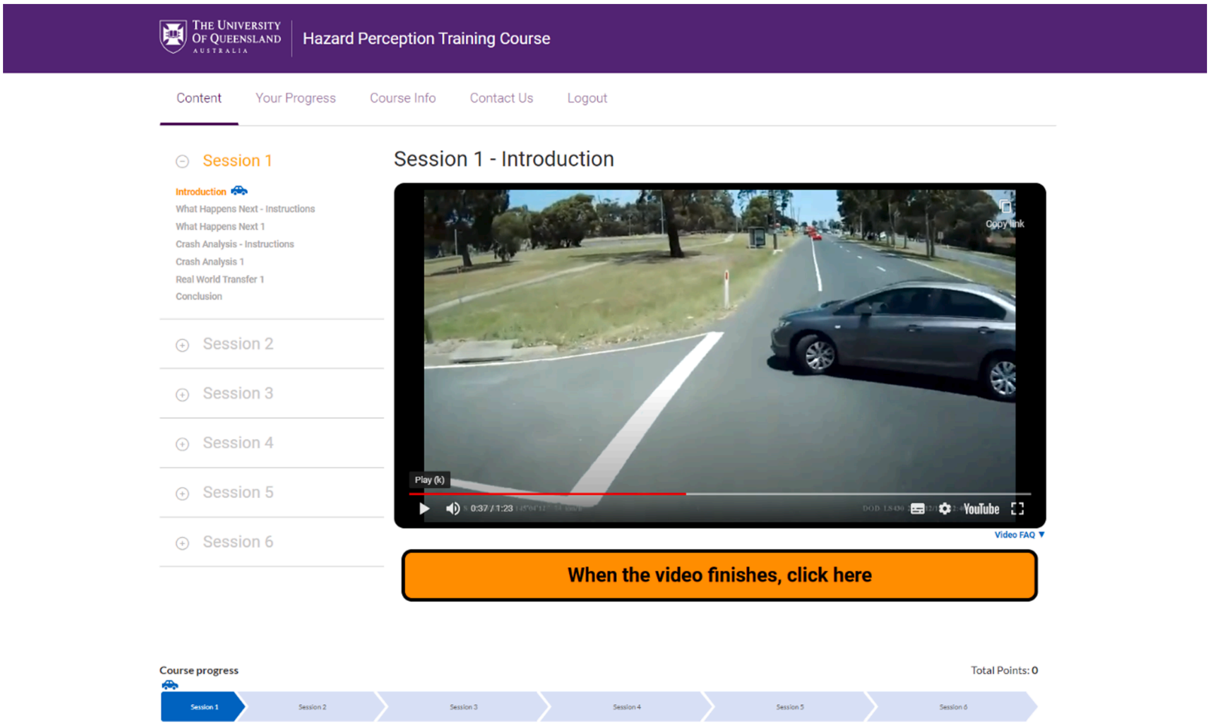

Fig. 1. Screenshot from the hazard perception training course.

potentially more motivated older drivers (Horswill, Falconer et al., 2015) was not found for novices (Wetton et al., 2013). Hence, persuading trainees of the value of improving their hazard perception might help improve retention of relevant learned skills, via improved engagement. This might also link to other deliberate practice strategies thought to be important to effective skill acquisition, such as goal-setting (Ericsson & Poole, 2016). For instance, drivers who have a clear goal of achieving a reduced crash likelihood might be more likely to retain what they have learned.

1.4. A hazard perception training course specifically designed to facilitate retention

We developed a hazard perception training course using these strategies (see Horswill, Hill, Silapurem et al., 2021, for full details), with the aim of creating an intervention that would result in long-term retention (in addition to other goals; see Section 2.2.1). The underlying approach of the course was to accelerate the development of drivers' mental representation of the traffic environment, by asking trainees to make judgements on videos of real traffic situations, with feedback. This was based on the well-established theoretical premise that drivers with superior hazard perception skill have a better situation awareness of the traffic environment, which they are able to use to predict dangerous situations earlier (Horswill & Hill, 2021).

Specifically, the course involved six half-hour sessions, each separated by a minimum of one week, to take advantage of the spacing effect and the benefits of sleep consolidation (see Fig. 1 for a screenshot from the course). The course included new versions of the two video-based exercise-types used by Wetton et al. (2013), but also included three other types of exercise (see Sections 2.2.1.1 to 2.2.1.5 for details). The different exercise types were interleaved within the sessions to take advantage of the contextual interference effect (see Section 2.2.1 for details of the course structure). The number of exercises completed was also considerably greater than in our previous hazard perception interventions (e.g., 8 for the combined intervention in Wetton et al., 2013). Finally, the course included messaging intended to persuade drivers that they would benefit from improving their hazard perception skill. This included using crash footage to illustrate how easily collisions can emerge from everyday driving situations, and how many of these potential crashes might be avoided with superior hazard perception skill. It was made clear to trainees they should set a goal of reducing their crash risk and that they could achieve this goal by engaging in the course.

1.5. Previous evaluations of the hazard perception training course

The short-term training effects of the course were evaluated in two previous studies, one with unsupervised novice drivers (Horswill, Hill, Silapurem et al., 2021) and the other with learner drivers who must drive with a licensed driver (Horswill, Hill, Bemi-Morrison et al., 2021). In both studies, significant and substantial training effects on hazard perception skill were found approximately one week after course completion (and the supervisors of the learner drivers were also found to have improved their hazard perception). However, these initial studies did not evaluate whether trainees' learning was retained beyond one week, and so the possibility remains that the training effects are short-lived.

1.6. The current study

The main goal of the present study is to evaluate whether the training effects of the course, as a whole, endure beyond the time span of these two previous evaluations, where previous work on other hazard perception training interventions indicates that this is not a foregone conclusion. The course was specifically designed using strategies to aid retention of learning (see Section 1.4) but, to date, has only been evaluated in the very short term. Evaluating medium-term retention is a crucial step in determining the efficacy of a hazard perception training intervention: if interventions are effective in the very short term but this effect disappears over longer time spans, then it raises questions about whether such a course is likely to be of practical value if deployed in the real world. To achieve this goal, we evaluated the effect of the training course on a validated measure of novice drivers' hazard perception skill both in the short term (defined as around one week post-training) and medium term (defined here as about two months post-training), using a randomized control design.

We hypothesize that the training course will improve trainees' hazard perception skill relative to controls, as in the previous studies. Given the evidence-based strategies that we adopted to promote skill retention (see Table 2 and Section 1.4), we also predict that this advantage will remain into the medium term, and that there will be no significant decrease in trainees' hazard perception skill between the short- and medium-term assessments.

For the present study, we chose to assess these training effects using a measure of hazard perception skill known as a *hazard prediction test*. In such a test, the driver views a series of video clips of traffic scenes (filmed from the driver's perspective). Each clip cuts to black before a traffic incident, at which point the driver is asked to predict what might happen next. Substantial validity evidence has been published for several hazard prediction tests, suggesting that this assessment method can yield scores that are a meaningful measure of hazard perception skill for research purposes (Crundall, 2016; Horswill et al., 2020; Jackson et al., 2009). Validity evidence for scores on the particular hazard prediction test used in the present study includes their empirical associations with crash involvement, which were found in two separate samples, and their ability to distinguish between novice and experienced drivers in both of those samples (Horswill et al., 2020). In addition, scores on the test were known to be sensitive to short-term improvements in hazard perception skill as a result of the training course under investigation (Horswill, Hill, Bemi-Morrison et al., 2021; Horswill, Hill, Silapurem et al., 2021), suggesting that, despite its brevity, the test is a good candidate for assessing potential skill retention or decay over a longer period.

Notably, many previous driver training interventions have been associated with counter-intuitive increases in crash involvement

(Christie, 2001). It has been proposed that if training leads to an increase in driver confidence, then this can fuel increases in risk-taking behaviour, which may reduce, negate, or even reverse any safety benefits of the training (Gregersen, 1996). Hence, the training course under evaluation was designed to avoid these increases by including training exercises previously found either to not inflate driver confidence (*commentary drive* and *what happens next* exercises, Horswill et al., 2013), or to significantly reduce it (*video review feedback* exercises, see Section 2.2.2.4, and Horswill et al., 2017).

Consistent with our aims in developing the course, completing the training has been found to have no impact on drivers' confidence (as indicated by self-ratings of driving skill) in the short term (Horswill, Hill, Silapurem et al., 2021). However, the possibility remains that increases in driver confidence could still occur over time (e.g., once drivers have had more opportunity to consolidate what they have learned). In the present study, we tested for this potential unintended consequence by evaluating drivers' self-ratings of their own skill in the medium term as well.

2. Method

2.1. Participants

Participants included in this study had to: (1) hold a current Australian driver's licence; (2) have passed an Australian on-road driving test; (3) be aged under 25 years; (4) typically drive at least weekly (in order to complete the *real world transfer* exercises, see Section 2.2.1.5); (5) have access to an email account and a mobile phone (for course registration); (6) have access to a PC with Windows 7 or later at home with a reliable internet connection (to allow them to complete the course); and (7) have not participated in other hazard perception research studies.

We recruited enough people to allow for a minimum of 84 participants at the medium-term assessment point (assuming substantial participant attrition due to the lengthy duration of the study), given 95% power, an alpha of 0.05, and assuming a large effect ($\eta^2 = 0.14$). Consequently, one hundred and fifty-two drivers were recruited. Of these, 92 responded to Facebook posts (these individuals were paid 200AUD; approximately 130EUR), and 60 took part via a first year psychology research participation pool (these individuals received course credit and were paid 100AUD; approximately 65EUR). All participants were randomly assigned to either a *trained group* or a *waitlist control group* (see Section 2.3 for details). Group assignment did not differ by recruitment method, X^2 (1, $N = 147$) = 0.33, $p = .567$ (and there was no significant interaction between group and recruitment method for either *hazard prediction test* scores or skill self-rating changes). One participant was excluded because they were assigned to the waitlist control group but erroneously completed the first session of the training course after the baseline assessment session. During the study, 4 participants in the *trained group* encountered difficulties accessing a working PC to complete the training course as intended due to COVID-19 movement restrictions in 2020. Their data were also excluded from the analysis.

Thirty-nine participants chose to withdraw before the short-term follow-up, leaving a final sample of 108 participants who contributed data to the study (see Section 2.3 for details of the appointment schedule and Table 1 for sample descriptives). This withdrawal rate was independent of study condition, X^2 (1, $N = 147$) = 3.00, $p = .083$. There were no significant differences at baseline between the two conditions for age (*Mann-Whitney* $U = 1409$, $p = .817$, *rank biserial correlation* = 0.025), years driving (*Mann-Whitney* $U = 1318$, $p = .433$, *rank biserial correlation* = 0.088), and kilometres driven per year (*Mann-Whitney* $U = 1243$, $p = .213$, *rank biserial correlation* = 0.140). Sex was also independent of study condition, X^2 (1, $N = 108$) < 0.01, $p = .960$. A further 4 participants withdrew from the study before the medium-term follow-up appointment, leaving 104 participants at that time-point. There was no significant difference between the trained and waitlist control groups in the number of individuals who completed the study and those who did not, X^2 (1, $N = 147$) = 2.84, $p = .092$.

Eighty-seven of the participants were recruited during March/April 2019 and the remaining 65 were recruited during March 2020. The 2020 sample completed the study during COVID-19 restrictions, which meant that they were tested online rather than in the

Table 1
Participant Characteristics ($n = 108$).

Variable		Trained group	Waitlist control group	Total
Age (years)	*M*	19.08	19.16	19.13
	SD	1.26	1.49	1.38
	Range	17–23	17–23	17–23
Sex	Female (%)	67.35	67.80	67.59
	Male (%)	32.65	32.20	32.41
Licence type	Provisional (%)[1]	97.96	98.31	98.15
	Unrestricted (%)[2]	2.04	1.69	1.85
Years since passing on-road driving test	*M* =	1.86	1.63	1.73
	SD =	1.11	0.80	0.95
	Range =	0.42–7.00	0.25–3.00	0.25–7.00
Kilometres driven per year	*M* =	13,159.88	11,168.78	12,072.15
	SD =	8,630.73	6,627.93	7,630.20
	Range =	1,591–58,000	165–30,000	165–58,000

[1] Provisional licence holders in Australia can drive unaccompanied (i.e., they are not learner drivers) but they are subject to certain driving restrictions.

[2] Unrestricted licence holders in Australia are not subject to the provisional driver restrictions.

laboratory (see Section 2.3). There were no significant group by year-of-testing interactions for changes in *hazard prediction test* scores or self-ratings of skill, indicating that the effects of the training on these variables did not differ significantly between the two years. Both group assignment, X^2 $(1, N = 147) = 0.36, p = .547$, and withdrawal rates (short-term: X^2 $(1, N = 147) = 1.91, p = .167$; medium-term: X^2 $(1, N = 147) = 1.52, p = .217$) were also independent of year-of-testing. Consequently, the samples from both years were merged for all substantive analyses.

2.2. Materials

2.2.1. Hazard perception training course

The current study evaluated a version of the hazard perception training course originally described by Horswill, Hill, Silapurem et al. (2021) and Horswill, Hill, Bemi-Morrison et al. (2021), which was modified with the addition of a further *real world transfer* exercise option (see Section 2.2.1.5 for details of this). In addition to the strategies designed to maximize enduring skill retention described earlier, the course also involved: (1) the use of over 100 video clips of real crashes; (2) evidence-based learning activities, to improve drivers' hazard perception skill (see Sections 2.2.1.1 to 2.2.1.4); (3) the inclusion of activities specifically designed to facilitate skill transfer to real driving (see Section 2.2.1.5); and (4) the application of "deliberate practice" principles and strategies (e.g., all exercises involved participants practicing hazard perception skill either on-road or with video clips of driving, see Sections 2.2.1.1 to 2.2.1.5).

The course was divided into six separate online sessions, with an enforced separation of no less than 7 days between each session during which participants were encouraged to complete a *real world transfer* exercise when driving (see section 2.2.1.5). The video-based elements of the course could be completed online on any networked PC running Windows 7 or later. A screenshot from the course can be seen in Fig. 1, Table 2 outlines the design rationales behind the course, and Fig. 2 shows the course structure. Descriptions of the various exercise types mentioned in Fig. 2 are provided below.

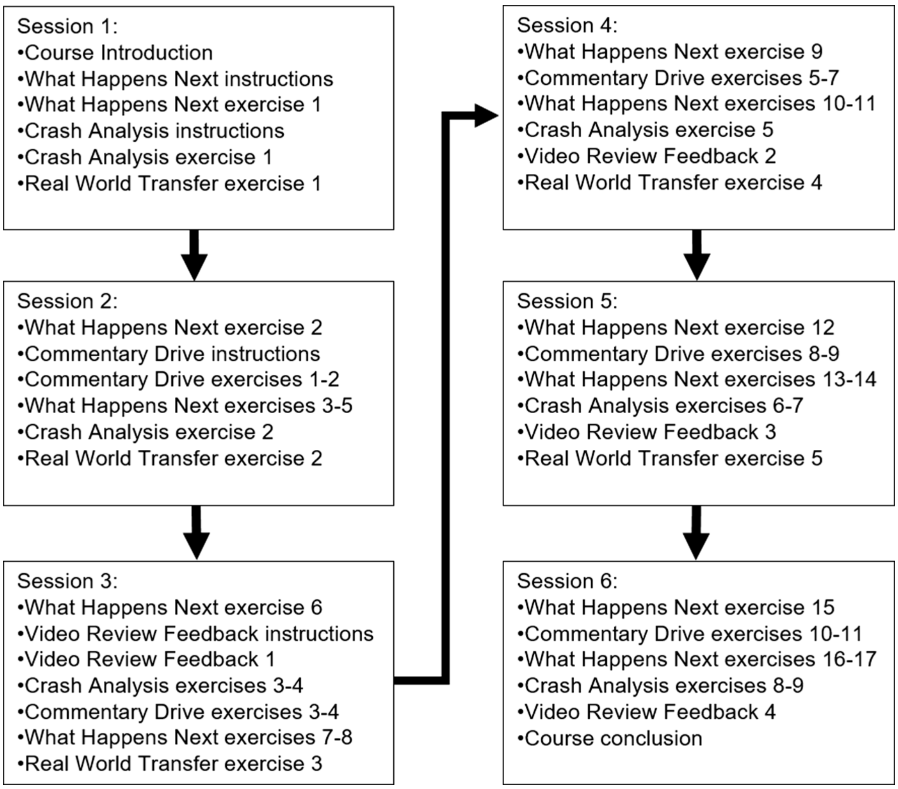

Fig. 2. The structure of the hazard perception training course (reproduced from Horswill, Hill, Silapurem et al., 2021).

Table 2
Summary of rationales behind the design of the new hazard perception training course (table adapted from Horswill, Hill, Silapurem et al., 2021). The first two rationales are proposed to be relevant to potentially improving retention of skills learned in the course. The third rationale relates to driver confidence.

Issues with previous driver training	Specifications for new interventions that address issues with previous training	Solutions used to meet specifications in the new course
Traditional training regimes do not typically take into account the forgetting curve.	Interventions should employ strategies known to minimize skill decay.	The course structure involves practice schedules known to reduce forgetting, via strategies such as distributed practice (e.g., spacing the training over 6 weekly sessions) and the interleaving of exercise-types. The format allows for convenient low-cost refresher sessions.
Driver training is considered more effective at improving safety when trainees are motivated to improve their safety and understand that this is the point of the training.	Interventions must be clear in communicating their goal to reduce crash risk (rather than just improving certain skills or enabling specific licensing tests to be passed) and may need to persuade trainees of the value of this goal.	The course frames learning goals explicitly in terms of avoiding crashes, and includes content designed to persuade trainees of the value of this goal (e.g., using footage of real crashes to illustrate how unpleasant these events are for those involved and how easily they can arise from everyday driving situations, see Section 1.4).
Some prior driver training has increased driver confidence, hence increasing risk-taking and crash risk.	Interventions must avoid triggering inappropriate increases in driver confidence and/or increases in risk-taking.	The course includes exercise-types previously found not to affect confidence ratings (e.g., commentary drive exercises, see Section 2.2.1.3).
Traditional driver skills training can be regarded as a far-transfer situation as drivers rarely experience the events that they are learning to avoid (i.e. crashes).	Interventions should create a near-transfer situation. Interventions should also consider strategies for increasing the likelihood of skill transfer.	The course provides high-frequency experience of real crashes using video clips. Multiple crash examples are used, drawing attention to underlying commonalities. On-road activities designed to facilitate transfer from the online course to real driving are included (e.g., drivers are asked to predict potential crash scenarios during real driving, see Section 2.2.1.5).
Traditional driver training focusses on lower-order skills that are not correlated with crash risk, such as vehicle handling. Training focussing on higher-order cognitive skills shows more promise.	Interventions should focus on improving higher-order skills known to be associated with crash involvement.	The course focuses on improving drivers' hazard perception skill (e.g., by including exercises previously shown to improve hazard perception, such as commentary drives, see Section 2.2.1.3).
Traditional driver training does not tend to systematically employ evidence-based strategies for the efficient attainment of expertise.	Interventions should employ evidence-based principles of efficient learning, where possible.	The course uses strategies associated with efficient learning, such as: goal-setting (e.g., trainees are asked to attempt to match expert drivers' commentary drive standards); deliberate practice (e.g., trainees are asked to generate their own commentary drives, both during video exercises and in real driving); provision of high quality and timely performance feedback (e.g., after generating each of their own commentary drives, trainees view an expert commentary on the same traffic situation to allow comparison); generative activities (e.g., drivers are asked to predict what will happen next in various situations, see Section 2.2.1.1); and the use of test-like exercises (e.g., trainees must identify traffic conflicts in videos while being timed, see Section 2.2.1.4).
Some driver training found to be effective involves high-cost resource-intensive instruction that may present barriers to mass roll-out and the use of refresher sessions.	Interventions should be scalable and low cost to facilitate mass roll-out and the use of refresher sessions.	The course is automated and can be run without an instructor. The course can be completed on any suitable internet-connected device.

2.2.1.1. What happens next exercises. The course included 17 *what happens next* exercises, spread across the 6 training sessions. In each exercise, the participant viewed a video clip of a traffic situation, which was filmed from the driver's perspective, while imagining that they were the driver. The clip cut to black before a traffic conflict (i.e., a situation in which the driver of the car with the camera would have to brake or steer to avoid hitting another road user) played out, in some cases resulting in a crash. When the clip cut to black, the participant was asked to describe any hazardous events that could occur within 5 s of the cut-point. Next, they viewed a video of an expert making predictions about the same clip, and were then shown how the incident in the clip actually resolved. Finally, they were shown examples of crashes that mapped onto other predictions made by the expert. The instruction video for this exercise-type can be viewed here: https://youtu.be/j6E0WkOCB4I.

2.2.1.2. Crash analysis exercises. Nine *crash analysis* exercises were spread across the 6 training sessions. In each exercise, the participant was shown a video clip of a real crash and then asked to analyse the incident, including describing what happened and why, as well as what those involved in the crash could do to minimize the chances of similar events reoccurring. Next, the participant viewed a video of an expert's analysis of the same crash. In some of these exercises, further examples of similar crashes, with commentary, were also shown to participants. The instruction video for this exercise can be viewed here: https://youtu.be/bv7G4RXY5C4.

2.2.1.3. Commentary drive exercises. Participants completed 11 *commentary drive* exercises, distributed across Sessions 2 to 6 of the course. In each exercise, the participant watched footage of a car journey (shot from the driver's perspective). Their task was to provide a running commentary describing what they were monitoring and what cues they could see that might signal potential hazards, as well as making predictions as to what might happen next. Then they viewed the same footage accompanied by an expert driver's commentary. Finally, they were shown a selection of crash clips to illustrate some of the comments made by the expert. The instruction video for this exercise-type can be viewed here: https://youtu.be/GtjFlV5M-jc.

2.2.1.4. Video review feedback exercises. A *video review feedback* exercise was included in each of Sessions 3 to 6 (4 exercises in total). Each exercise included 4 or 5 video clips filmed from the driver's perspective, which contained traffic conflicts (i.e., events in which the camera car was on course to hit another road user at some point). For each clip, participants were asked to click on the other road user (s) involved as soon as they predicted a traffic conflict. The clip was then replayed with superimposed markers indicating the time and location of (1) the participant's response, (2) the point where the average driver would predict the conflict, and (3) the point where an expert driver would predict the conflict (average and expert driver data were derived from previous work, see Horswill et al., 2017, for details). The same data were then displayed as response times on a bar graph. The replay and graph provided performance feedback to allow the participant to see how they compared with other drivers. After completing all of the clips in an exercise, participants were shown video clips of crashes associated with traffic conflicts similar to those in the clips. The instruction video for this exercise-type can be viewed here: https://youtu.be/avdZ6dKdHwU.

2.2.1.5. Real world transfer exercises. At the end of Sessions 1 to 5, the participant was asked to select an activity to complete during real driving in the interval before the following session. The aim of this exercise was to facilitate transfer between what they had learned during the computer-based components of the course and actual traffic situations. Each activity was linked to an element of the

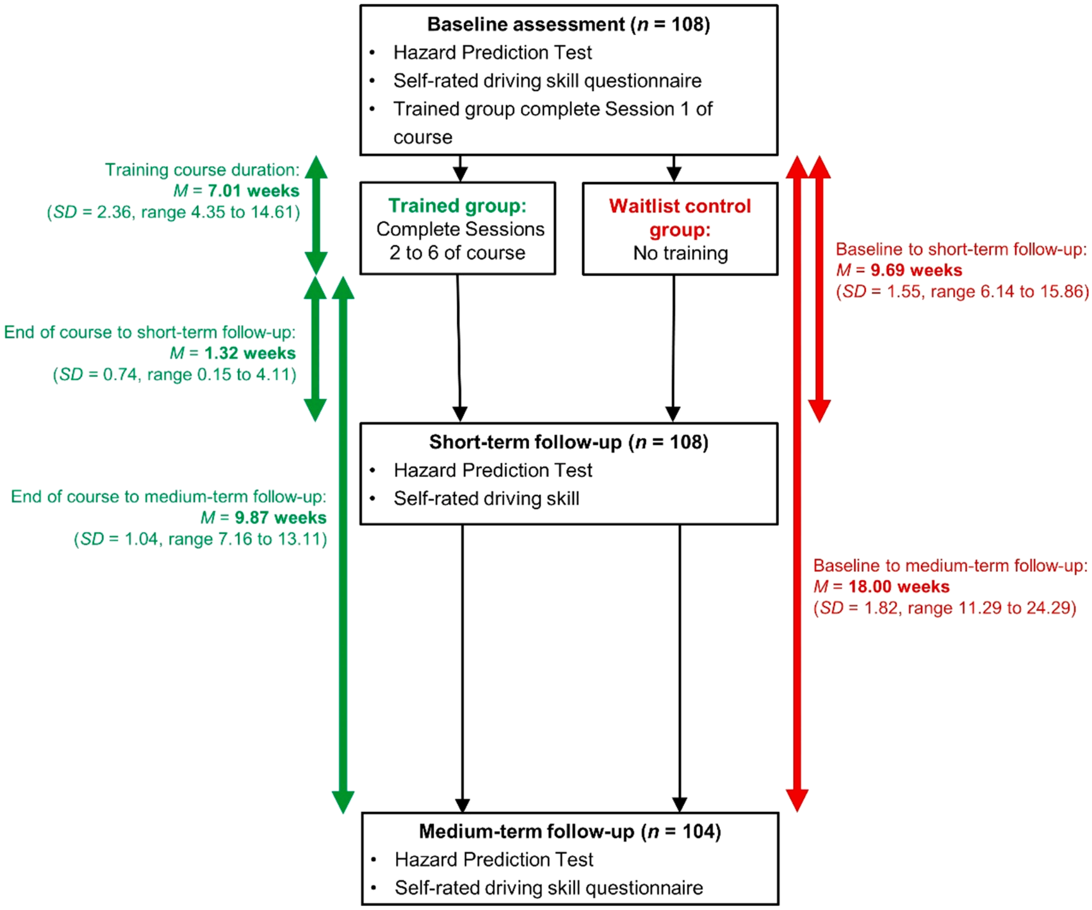

Fig. 3. Study design.

video-based exercises. There were two options available in Session 1, with new options introduced each session, to make a final total of six. Options included: “At various points during my drive, I will consider the following question: If a crash was going to occur within the next five seconds, what could it be?”; and “I will make a special effort to monitor and adjust my following distances, based on my assessment of the current level of risk”. One new option was introduced in the version of the course used in the present study (which was the only change made to the course). In this option, the participant’s task was to “make a special effort to adjust my speed, based on my assessment of the current level of risk”.

After selecting their *real world transfer* activity, the participant was asked to choose a day and time to receive a reminder text message (e.g., just before an anticipated journey that week). At the start of the next computer-based training session, the participant completed a survey in which they were asked whether they engaged in their chosen activity. 89.80% of trainees reported completing all five *real world transfer* activities, 8.16% reported missing one, and 1 individual reported missing two. No one reported missing more than two activities.

2.2.2. Hazard prediction test

Participants completed a validated *hazard prediction test* (Horswill et al., 2020) during each of the three laboratory appointments (see Section 2.3 and Fig. 3). The test was described to participants as a “What Happens Next” assessment, and no mention was made of the term “hazard” in the test instructions. This was because test-takers can vary substantially in how “hazardous” they judge a particular driving scenario to be (Wetton, Hill, & Horswill, 2011). For example, researchers have found cross-cultural differences in the interpretation of what constitutes a traffic hazard (Ventsislavova et al., 2019). Hence, instructing drivers to predict “hazards” potentially conflates the measurement of anticipatory skill (i.e., *hazard perception*) with test-takers’ subjective judgements as to how dangerous an event must be before it is regarded as hazardous (i.e., *risk perception*).

The test featured six clips filmed from the driver’s perspective. The participant was invited to imagine that they were driving the camera car. Each clip depicted the precursors to several potential traffic conflicts (traffic conflicts were defined as a situation in which the participant, as driver of the camera car, would need to slow down or change course to prevent a crash).

The clip cut to black after the appearance of one or more predictive cues, but before any traffic conflicts had a chance to play out. The participant’s task was to describe as many predictions as they could, out loud, regarding what might happen next. Their responses were audio recorded and an experienced assessor, who was blind to the participants’ condition, scored them. To determine how many plausible traffic conflicts participants predicted, the assessor made a judgement as to whether each of the participant’s predictions was a sufficient match to one of the items on a pre-determined list of predictions derived from an expert panel (see Horswill et al., 2020, for details). The assessor assigned participants one point for every matching prediction. The participant’s test score was calculated as the average number of matching predictions made per clip. These scores have been found, in two separate samples, to correlate with self-reported crash involvement and to discriminate between high-risk (novice) and lower-risk (experienced) driver groups (Horswill et al., 2020). The internal consistency of the test in the current study (baseline assessment) was acceptable (*Cronbach’s* $\alpha = 0.72$).

To examine the transferability of the training to previously unseen clips, four new clips were added for the short-term assessment (*Cronbach’s* $\alpha = 0.81$) and a further four were added for the medium-term assessment (*Cronbach’s* $\alpha = 0.70$). These clips were shown after the six original clips at each time-point, and analysed separately.

2.2.3. Self-rated driving skill questionnaire

Participants completed a questionnaire in which they rated their own driving skill, both overall and for 15 different skill components, as compared with the median driver (Horswill et al., 2013; Horswill et al., 2004). Six of the measured skills were related to *hazard perception* (e.g., “awareness/anticipation of other road users’ behaviour”) and four were related to *vehicle control* (e.g., “controlled emergency stops”). The *vehicle control* items were intended as a control measure, where the hazard perception training course might be expected to affect hazard perception skill ratings but not vehicle control ratings.

Note that the original questionnaire had 18 items, but 2 items were excluded because they referred to manual gear use (“hill starts” and “appropriate use of gears”), which many participants in the present sample would have been unfamiliar with. There was one item previously classified as *vehicle control* (Horswill et al., 2004), which was also excluded in this study (“adapting to conditions”) because the concept of adapting one’s driving to suit the conditions was explicitly linked to hazard perception during the training course.

Participants responded to all items on an 11-point scale (1 = “Bottom 10% of Brisbane drivers”; 6 = “Typical Brisbane driver – 50% are more skilful, 50% are less skilful”; 11 = “Top 10% of Brisbane drivers”). For each participant, sub-scale scores were calculated by averaging over the relevant items, and a composite measure of *overall driving skill* was determined by averaging self-ratings for the 15 specific skill components and the single overall skill item. The internal consistency of the scales in the present study was acceptable (hazard sub-scale: *Cronbach’s* $\alpha = 0.82$; vehicle control sub-scale: *Cronbach’s* $\alpha = 0.82$; overall scale: *Cronbach’s* $\alpha = 0.92$).

2.3. Procedure

Fig. 3 shows the design of the randomized control trial. Participants completed the outcomes measures (the *hazard prediction test* and the self-rated driving skill questionnaire) at three different time points (baseline, short-term post-course, and medium-term post-course). For participants recruited in 2019, all assessments were conducted face-to-face in a university laboratory. For participants recruited during 2020, all assessments were conducted online due to restrictions associated with the COVID-19 pandemic. Online assessments involved the use of video conferencing to monitor participants as they completed online versions of the three tests.

After completing the baseline assessment, participants in the waitlist control group were told that they would be contacted to arrange a follow-up appointment. The trained group commenced the hazard perception training course immediately, completing

Session 1 under supervision (either in the laboratory or monitored via video-conferencing software). The trained group were then asked to schedule dates on which they would complete the remaining 5 training sessions (unsupervised), leaving at least 1 week between each session.

Researchers monitored the progress of participants in the trained group, sending reminder messages before each session and follow-up messages if they missed a scheduled session, asking them to reschedule.

The short-term follow-up assessment was scheduled for approximately one week after the trained group had completed the training course, and the medium-term follow-up was scheduled for approximately two months after the end of the course (see Fig. 3 for descriptives). At the end of the medium-term follow-up assessment, participants in the waitlist control group were given access to the course to complete in their own time. Note that other driving-related measures were also taken using this sample of participants, but these were not relevant to the present study.

2.4. Data analysis

The effects of the training course on hazard perception skill were evaluated using a series of ANCOVAs, with a post-training *hazard prediction test* score (either short-term or medium-term, as appropriate) as the dependent variable. In all analyses, *trained* versus *control* assignment was the independent variable, and *baseline hazard prediction test score* was the covariate.

Potential training decay effects between the short and medium-term assessments were assessed using paired-sample t-tests on the *hazard prediction test* scores. The effects of the training course on self-ratings of driving were evaluated using independent sample t-tests (independent variable: group; dependent variable: change in rating between baseline and both post-training assessment points).

3. Results

3.1. Effects of the training course on hazard perception skill: Trained participants vs. controls

One hundred and eight participants provided data for the short-term follow-up (though *hazard prediction test* data for 8 of these participants was lost due to technical issues). One hundred and four participants provided data for the medium-term follow-up (no missing data).

Using Horswill et al.'s (2020) original 6-clip *hazard prediction test* (where the same clips were used at both baseline and post-training, noting that participants received no feedback regarding their performance), a significant training effect (i.e. a trained vs. untrained advantage, controlling for baseline performance) was found at both the short-term follow-up, $F(1,97) = 28.9$, $p < .001$, $\eta^2 =$

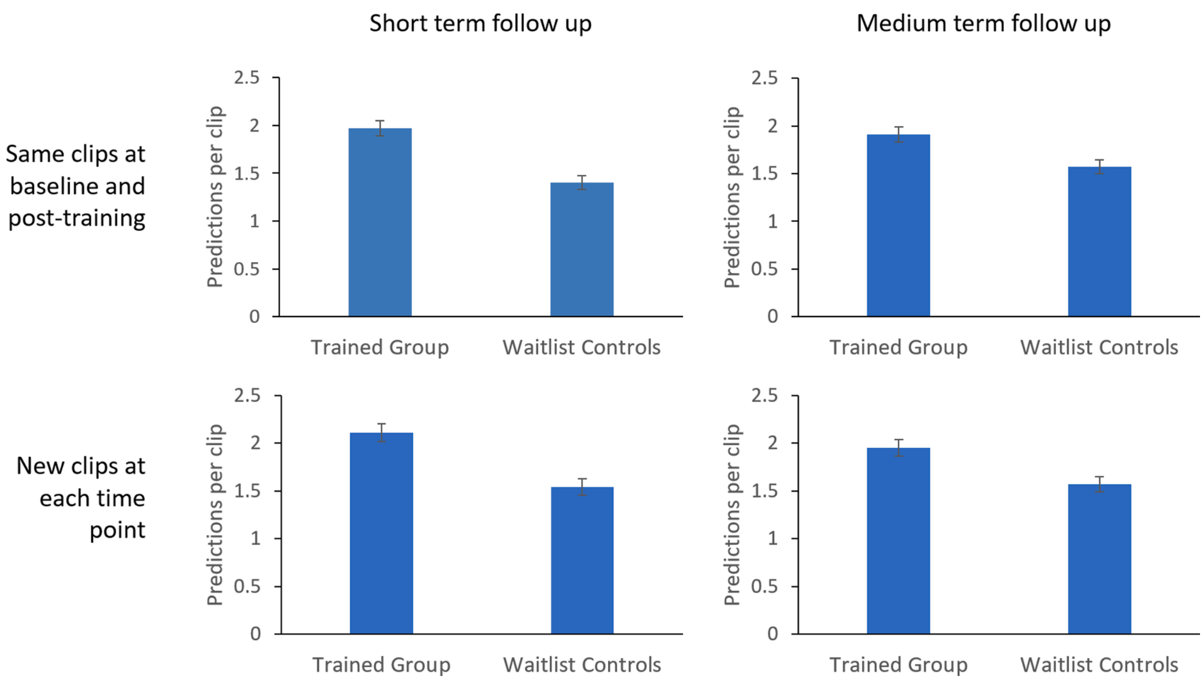

Fig. 4. *Hazard prediction test* scores by group (estimated marginal means, covarying for baseline performance; error bars show standard errors of the mean) at short-term follow-up (left column) and medium-term follow-up (right column), for the same clips at baseline and post-training (top row), and for new clips introduced at each post-training time point (bottom row). Note that the maximum possible number of matched predictions (averaged across clips) was 4.33 for the 6-clip test administered at baseline and post-training, and 4.75 and 4.50 for the two follow up tests involving new clips.

Table 3
Raw means (SDs) of hazard prediction test scores (mean matched predictions per clip) by group across the three assessment points.

	Baseline assessment	Short-term assessment	Medium-term assessment
Trained group	1.23 (0.47)	1.92 (0.70)	1.86 (0.67)
Waitlist control group	1.36 (0.52)	1.44 (0.56)	1.62 (0.61)

0.168, $\omega^2 = 0.161$, and the medium-term follow-up, $F(1,101) = 10.1$, $p = .002$, $\eta^2 = 0.066$, $\omega^2 = 0.059$ (see Fig. 4 for estimated marginal means and Table 3 for raw means). This indicates that trained participants were able to predict more traffic conflicts than untrained participants at both post-training assessment points.

In order to evaluate the generalisability of the training effect, further ANCOVAs were conducted using only data from the previously unseen *hazard prediction test* clips presented during the post-training assessments. As before, significant effects were found both at the short-term follow-up, $F(1,97) = 19.0$, $p < .001$, $\eta^2 = 0.119$, $\omega^2 = 0.112$, and the medium-term follow-up, $F(1,101) = 9.94$, $p = .002$, $\eta^2 = 0.072$, $\omega^2 = 0.064$ (see Fig. 4 for estimated marginal means). Note that scores for unseen clips violated normality at the short-term follow-up ($p = .006$), but data from a square root transformation to remove this violation yielded a nearly identical result, $F(1,97) = 17.43$, $p < .001$, $\eta^2 = 0.108$, $\omega^2 = 0.102$. That is, the ability of trained participants to predict more traffic conflicts than untrained participants extended to situations they had not previously encountered.

Given the individual differences in training course duration, we inspected the correlation between training course duration and the change in the hazard prediction test score (6-clip version) between baseline and short-term follow up. The correlation was not significant, $r(44) = -0.06$, $p = .714$, indicating that variation in the course duration was not associated with training effectiveness.

3.2. Decay of training: short-term vs. medium-term hazard perception performance

The *hazard prediction test* scores of the trained drivers did not significantly decrease between the short- and medium-term assessments, $t(43) = 0.80$, $p = .428$, *Cohen's d* = 0.121 (original six-clip test). This suggests that the original effect of training was not diminished by the passage of time (around 2 months). However, there was also a significant *increase* in *hazard prediction test* score for the waitlist control group from the short term to the medium term, $t(51) = -3.07$, $p = .003$, *Cohen's d* = -0.426. This increase could not be accounted for by individual differences in the duration between the short and medium-term assessments, as the correlation between duration and hazard prediction test scores (6-item test) was not significant, $r(94) = -0.10$, $p = .327$. Possible explanations for the improvement of the control group (noting that neither group received training between the short-term and medium-term assessments) are provided in the discussion. Note that there was a violation of normality in each of these analyses, but non-parametric tests yielded the same pattern of results (trained group, short-to-medium term: *Wilcoxon W* = 407.00, $p = .601$, *rank biserial correlation* = -0.099; control group, short-to-medium term: *Wilcoxon W* = 272.00, $p = .009$, *rank biserial correlation* = -0.451).

3.3. Effects of the training course on self-ratings of driving skill: Trained participants vs. controls

There was no training effect on any of the self-ratings of driving skill, either at short-term follow-up (*hazard perception*: $t(106) = -0.79$, $p = .433$, *Cohen's d* = -0.15; *vehicle control*: $t(106) = 1.07$, $p = .289$, *Cohen's d* = 0.21; *overall driving skill*: $t(106) = 0.20$, $p = .845$, *Cohen's d* = 0.04), or medium-term follow-up (*hazard perception*: $t(102) = -0.11$, $p = .915$, *Cohen's d* = -0.02; *vehicle control*: $t(102) = 0.30$, $p = .765$, *Cohen's d* = 0.06; *overall driving skill*: $t(102) = -0.12$, $p = .90$, *Cohen's d* = -0.02). Note that there was a violation of normality for the short-term effects, but non-parametric tests yielded the same pattern of results (*hazard perception*: *Mann-Whitney U* = 1226.50, $p = .177$, *rank biserial correlation* = 0.15; *vehicle control*: *Mann-Whitney U* = 1329.5, $p = .474$, *rank biserial correlation* = 0.08; *overall driving skill*: *Mann-Whitney U* = 1437.50, $p = .963$, *rank biserial correlation* = 0.01).

The duration of the training course did not correlate with self-rated changes between baseline and the short-term assessment, $r(47) = -0.06$, $p = .663$, indicating that this factor was not associated with perceived course outcomes.

4. Discussion

4.1. Key study findings and their implications

This study describes a randomized control trial that was conducted to evaluate the effectiveness of a hazard perception training course for improving novice drivers' hazard perception skill in both the short term (*M* = 9.24 days, or 1.32 weeks; *range:* 1.05 to 28.78 days) and medium term (*M* = 9.87 weeks, or 2.27 months; *range:* 50.09 to 91.78 days). Significant and substantial training effects were found for the trained group relative to a waitlist control group in scores on a validated *hazard prediction test* in both the short-term and medium-term follow-up assessments. These improvements were found for traffic situations seen by all participants at baseline, and also for novel traffic situations introduced at each follow-up assessment. Furthermore, there was no evidence of skill decay among the trained group, as their scores did not significantly decrease between the short-term and medium-term follow-ups.

These results suggest that the new course has succeeded in achieving retention of novice drivers' hazard perception learning over this time frame, indicating that the course has the capacity to improve hazard perception skill for at least two months. Nevertheless, it is possible that this skill will decay over a longer time frame, which is something that needs to be investigated in future research. However, even if this is the case, the nature of the intervention (online, automated, with relatively low costs to make the materials available to drivers) means that booster sessions with appropriate spacing (e.g., every few months) could potentially be used.

4.2. Potential mechanisms facilitating the successful retention

These findings contrast with the skill decay found in some prior studies in the weeks and months following hazard perception training (e.g., Wetton et al., 2013; Reyes and O'Neal, 2020). While direct comparisons between the different studies are not possible (e.g., different assessment timing, different dependent measures, different designs, different samples), we can speculate on possible reasons for the retention of learning found in the present study. These might include: (1) the use of training strategies known to increase retention (e.g., the spacing effect and high contextual interference); (2) the relatively large number of exercises and correspondingly greater course duration (46 exercises over about 3 h); (3) greater trainee engagement with the course (e.g., due to the use of real crash examples); and (4) greater motivation to learn among trainees, due to course content devoted to persuading them of the value of the training.

The present findings do not provide insight into which (if any) of these potential sources of variation are responsible for the retention of training in this study. Future work to determine the conditions required for retention in this context would be useful. For instance, if it was found that similar retention could be obtained with a shorter training course, then this might provide further benefits (e.g., a shorter course might lead to greater training uptake and completion).

4.3. Why waitlist control participants improved between the short and medium term assessments

One unexpected finding was that the waitlist control group significantly increased their *hazard prediction test* scores on Horswill et al.'s (2020) original 6-clip *hazard prediction test* between the short- and medium-term assessments, despite both groups experiencing the same treatment during this time (i.e., no intervention). This led to a reduced, though still significant, difference between the two groups at the medium-term assessment (where the effect size for the between-groups difference declined from large to medium). There are several possible explanations for this finding. First and foremost, given the extended time frame involved, the waitlist control group could potentially have benefitted from maturation of their hazard perception skill through the driving experience they accrued in the time between the short-term and medium-term follow-up assessments (given that hazard perception skill improves with experience, Horswill & Hill, 2021). In contrast, those in the trained group may have been rendered precocious by the training, effectively skipping this maturation period by fast-tracking their skill development.

Other explanations are possible. For instance, assuming that they recognised the traffic clips from the baseline assessment, we can speculate that the trained group probably had insight that the course had substantially improved the number of predictions that they made during the short-term follow-up assessment, and had no expectation of further improvements at the medium-term follow-up assessment since they had received no further training in the interim. Hence, in the medium-term follow-up test, they may have been satisfied with their performance once they had identified a similar number of hazards to what they identified in the short-term follow-up test. In contrast, the control group might have had some insight that they were performing at a low level on this test (despite receiving no feedback on their performance). This may have manifested in either genuine attempts to perform better, leading them to deduce some of the more generic predictions (i.e., outcomes that are less reliant on specific training to predict), or increased guessing. Alternatively, it is possible that the much greater number of video clips experienced by participants in the trained group as part of the course might have made them less likely to remember details of the 6 original clips from the *hazard prediction test*. The control group, however, only saw the clips used in the *hazard prediction test,* and hence might have been better able to remember and reflect on those specific clips between assessments.

4.4. The medium term effects of the training course of drivers' self-ratings of skill

Despite improving their hazard perception skill, completing the training course did not affect participants' self-ratings of their own driving skills (i.e., *hazard perception skill*, *vehicle control skill*, and *overall driving skill*) in either the short or medium term. These findings are consistent with prior research on the short-term effects of the training course (Horswill, Hill, Silapurem et al., 2021), and extend them to the medium term. Nevertheless, it was possible that the course *could* have yielded reductions in self-ratings in the present study, given that some of its constituent elements have been associated with decreased self-ratings in the past, namely the *video review feedback* exercises (Horswill et al., 2017). More crucially, however, both the present findings and those of Horswill, Hill, Silapurem et al. (2021) suggest that the course does not appear to lead to increased driver confidence, which has been linked to negative learning and increased crash risk in the past (Gregersen, 1996).

4.5. Study limitations and future research

A key limitation of the current study is that it provides little insight into which of the elements of the training course were responsible for the medium-term retention (see Section 4.2). Further work, involving different versions of the course, in which key elements are added or removed, would be needed to tease out the underlying strategies that facilitated the medium-term retention, which has not been found with other interventions.

Another limitation is that it is possible that the training course improved test scores without improving real driving behaviour, despite the prior evidence linking scores on the *hazard prediction test* that we used to crash involvement (Horswill et al., 2020). However, while examining the transfer of learning from the course to real driving was beyond the scope of the data presented in this paper, we have examined it elsewhere (Horswill et al., 2022). Specifically, the course was found to reduce heavy braking, speeding, and over-revving rates during everyday driving (though note that, unlike the present study, the on-road study did not evaluate medium term retention).

Other useful avenues for future research include evaluating the course over longer time frames, with different driver groups (e.g., drivers aged 65 and over), and examining whether the course directly impacts drivers' crash involvement. Individual differences in training effectiveness could also be examined. For example, there is some data suggesting that hazard perception training benefits differ according to driver personality factors (Zhang et al., 2018). This could mean that the most effective approach would including having different intervention protocols for different drivers. For instance, some individuals may be able to gain the full training benefits from a shorter version of the course without supervision, while others may need one-on-one mentoring in a more controlled environment over a longer training period. In addition, further research to determine the relative contributions of the different elements of the course would be of value. For instance, it is presently unclear the extent to which new elements, such as the ***real world transfer*** activities or the inclusion of real crash footage, contribute to the overall training effect. Another open question is whether the automated online nature of the course creates disadvantages compared with if it had been deployed face-to-face, potentially with an instructor. While the former has clear cost and convenience benefits, it could be that there are benefits from a face-to-face instructor-led alternative, such as increased engagement and the ability to tailor the training for individuals.

4.6. Practicalities regarding potential deployment of the hazard perception training course

In terms of enabling large scale deployment of the hazard perception training course, further investigation of how to persuade drivers to complete the course is warranted. In the present study, considerable financial compensation (or course credit equivalents) was provided, but this may not be practical when rolling out the course at scale. Other options to increase the likelihood of course completion might include restructuring the course to fit into existing driver training (e.g., trainees do brief sessions on a mobile device at the beginning of their traditional driving lessons) or leveraging influential others (e.g., parent-supervisors) to encourage course completion at the learner stage (Horswill, Hill, Bemi-Morrison et al., 2021). In jurisdictions where drivers are required to obtain a minimum score on a hazard perception test as a requirement of obtaining their unrestricted driver licence (e.g., the UK and Australia), one option would be to make hazard perception courses available to individuals about to attempt the test (e.g., as part of online test preparation materials). The incentive to complete the training in this situation would be that trained drivers would be more likely to pass the test and hence be able to progress towards obtaining their unrestricted drivers' licence.

4.7. Technical aspects of the study: face-to-face versus online testing

One notable technical aspect of the current study was the shift from face-to-face to online testing partway through due to COVID restrictions. While it was not possible to draw empirical conclusions about this shift because participants were not randomly assigned to the two conditions, we can speculate about the likely impact. Differences with the online experience would include: (1) the equipment was not standardized, as online participants used their own computers rather than lab computers; (2) the environment was not standardized (i.e., each online participants completed the assessments in a different room); and (3) it is possible that there were distractions that were not evident to researchers monitoring participants remotely that would have been noticeable if researchers had been face-to-face. The expected effect of these factors would be to potentially add noise into the data, assuming any of the factors had a meaningful influence on task performance (e.g., via increased distraction). This could, in principle, lead to online testing yielding smaller effect sizes as a result (hence requiring larger sample sizes to maintain the same statistical power).

4.8. Conclusions

In conclusion, the results presented in this paper suggest that the hazard perception training course under evaluation can improve hazard perception skill (measured using scores on a test known to be associated with crash risk), and that improved performance is retained in the medium term. Importantly, the performance improvements include generalization to novel traffic situations. While further data is required to demonstrate transfer to real driving and crash risk, the findings nonetheless support the proposal that this type of training intervention has the potential to reduce crash risk on a large scale.

Author Note

The authors thank Paul Jackson for programming the software used in this project. We also thank the following people for their help with participant testing: Jack Chapman, Thomas Cross, Sage Cullen, Mingqian Deng, Mathilde Gargan, Jian Guo, Samuel Halpin, Jessica Hellwig, Yat Fung Lai, Ho Yin Lam, Lauren Moody-Stone, Sheung Yan Ng, Melanie Parkes, Emily Porter, Adam Price, Nathan Simpson, Nonie Siu, Nicky Wildermuth, Chung Pui Grace Lam, William Stobo, Rachel McKinnon, Chynna Galon, Guillermo Monroy, Nicole Delmo, Brendan Thomson, Winta Nwatu, Taneil Davidson, Jana Staby, Roman Rao, Ji Yeon Kim, Georgina Baker, Madison Neve, Joseph Lim, Wai Shan Cheung, Gisele Faria, Mikayla Hancock, Tallara Drew-Hazou, Tao Li, Rey Yago, Nathan Chand, Maharshi Patel, and Shanice Forsythe.

Funding

The ACT Road Safety Fund Grants Program (RSG0082018) provided funding for this study.

CRediT authorship contribution statement

Mark S. Horswill: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing, Funding acquisition. **Andrew Hill:** Conceptualization, Methodology, Resources, Software, Visualization, Writing – review & editing, Funding acquisition. **Lisa Buckley:** Conceptualization, Methodology, Supervision, Project administration, Resources, Writing – review & editing, Funding acquisition. **Genevieve Kieseker:** Methodology, Investigation, Resources, Data curation, Project administration, Supervision, Writing – review & editing. **Francine Elrose:** Methodology, Investigation, Resources, Data curation, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

References

Ahmadi, N., Katrahmani, A., & Romoser, M. R. (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, 1965-1969. Sage.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2020). *Memory* (3rd ed.). Routledge.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General, 120*(1), 20. https://doi.org/10.1037/0096-3445.120.1.20

Boufous, S., Ivers, R., Senserrick, T., & Stevenson, M. (2011). Attempts at the practical on-road driving test and the hazard perception test and the risk of traffic crashes in young drivers. *Traffic Injury Prevention, 12*(5), 475–482. https://doi.org/10.1080/15389588.2011.591856

Broadbent, D. P., Causer, J., Williams, A. M., & Ford, P. R. (2017). The role of error processing in the contextual interference effect during the training of perceptual-cognitive skills. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1329–1342. https://doi.org/10.1037/xhp0000375

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354. http://psycnet.apa.org/journals/bul/132/3/354/.

Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained drivers. *Transportation Research Part F, 5*, 157–167. https://doi.org/10.1016/S1369-8478(02)00014-1

Cheng, A. S. K., Ng, T. C. K., & Lee, H. C. (2011). A comparison of the hazard perception ability of accident-involved and accident-free motorcycle riders. *Accident Analysis and Prevention, 43*(4), 1464–1471. https://doi.org/10.1016/j.aap.2011.02.024

Christie, R. (2001). *The effectiveness of driver training as a road safety measure: A review of the literature* (Report No. 01/03). Royal Automobile Club of Victoria. https://s3-ap-southeast-2.amazonaws.com/cdn-nrspp/wp-content/uploads/sites/4/2017/07/20094955/The-Effectiveness-Of-Driver-Training-As-A-Road-Safety-Measure-A-Review-Of-The-Literature.pdf.

Chihuri, S., et al. (2016). Driving cessation and health outcomes in older adults. *Journal of the American Geriatrics Society, 64*(2), 332–341. https://doi.org/10.1111/jgs.13931

Crundall, D. (2016). Hazard prediction discriminates between novice and experienced drivers. *Accident Analysis and Prevention, 86*, 47–58. https://doi.org/10.1016/j.aap.2015.10.006

Darby, P., Murray, W., & Raeside, R. (2009). Applying online fleet driver assessment to help identify, target and reduce occupational road safety risks. *Safety Science, 47*(3), 436–442. https://doi.org/10.1016/j.ssci.2008.05.004

Deery, H. A., & Love, A. W. (1996). The effect of a moderate dose of alcohol on the hazard perception profile of young drink-drivers. *Addiction, 91*, 815–827. https://doi.org/10.1111/j.1360-0443.1996.tb03577.x

Ericsson, A., & Poole, R. (2016). *Peak: Secrets from the new science of expertise.* The Bodley Head.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Glendon, A. I. (2014). An approach to novice driver training. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 64*(3), 111–122. https://doi.org/10.1016/j.erap.2014.04.003

Gregersen, N. P. (1996). Young drivers' overestimation of their own skill - An experiment on the relation between training strategy and skill. *Accident Analysis and Prevention, 28*(2), 243–250. https://doi.org/10.1016/0001-4575(95)00066-6

Hill, A., Horswill, M. S., Whiting, J., & Watson, M. O. (2019). Computer-based hazard perception test scores are associated with the frequency of heavy braking in everyday driving. *Accident Analysis and Prevention, 122*, 207–214. https://doi.org/10.1016/j.aap.2018.08.030

Horswill, M. S. (2016a). Hazard perception in driving. *Current Directions in Psychological Science, 25*(6), 425–430. https://doi.org/10.1177/0963721416663186

Horswill, M. S. (2016b). Improving fitness to drive: The case for hazard perception training. *Australian Psychologist, 51*, 173–181. https://doi.org/10.1111/ap.12132

Horswill, M. S., Anstey, K. J., Hatherly, C. G., & Wood, J. (2010). The crash involvement of older drivers is associated with their hazard perception latencies. *Journal of the International Neuropsychological Society, 16*(5), 939–944. https://doi.org/10.1017/S135561771000055X

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

Horswill, M. S., Garth, M., Hill, A., & Watson, M. O. (2017). The effect of performance feedback on drivers' hazard perception ability and self-ratings. *Accident Analysis and Prevention, 101*, 135–142. https://doi.org/10.1016/j.aap.2017.02.009

Horswill, M. S., & Hill, A. (2021). Drivers' hazard perception skill. In R. Vickerman (Ed.), *International Encyclopedia of Transportation* (Vol. 7, pp. 151-157). Elsevier. https://doi.org/10.1016/B978-0-08-102671-7.10678-5.

Horswill, M. S., Hill, A., Bemi-Morrison, N., & Watson, M. O. (2021). Learner drivers (and their parent-supervisors) benefit from an online hazard perception course incorporating evidence-based training strategies and extensive crash footage. *Accident Analysis and Prevention, 161*, 106340. https://doi.org/https://doi.org/10.1016/j.aap.2021.106340.

Horswill, M. S., Hill, A., Buckley, L., Kieseker, G., & Elrose, F. (2022). An online hazard perception training course reduces heavy braking, speeding, and over-revving rates during everyday driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 87*, 54–68. https://doi.org/10.1016/j.trf.2022.03.018

Horswill, M. S., Hill, A., & Jackson, T. (2020). Scores on a new hazard prediction test are associated with both driver experience and crash involvement. *Transportation Research Part F: Traffic Psychology and Behaviour, 71*, 98–109. https://doi.org/10.1016/j.trf.2020.03.016

Horswill, M. S., Hill, A., Silapurem, L., & Watson, M. O. (2021). A thousand years of crash experience in three hours: An online hazard perception training course for drivers. *Accident Analysis and Prevention, 152*, Article 105969. https://doi.org/10.1016/j.aap.2020.105969

Horswill, M. S., Hill, A., & Wetton, M. (2015). Can a video-based hazard perception test used for driver licensing predict crash involvement? *Accident Analysis and Prevention, 82*, 213–219. https://doi.org/10.1016/j.aap.2015.05.019

Horswill, M. S., & McKenna, F. P. (1999). The effect of interference on dynamic risk-taking judgments. *British Journal of Psychology, 90*(5), 189–199. https://doi.org/10.1348/000712699161341

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury, & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and Application* (pp. 155–175). Ashgate.

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

Horswill, M. S., Waylen, A. E., & Tofield, M. I. (2004). Drivers' ratings of different components of their own driving skill: A greater illusion of superiority for skills that relate to accident involvement. *Journal of Applied Social Psychology, 34*(1), 177–195. https://doi.org/10.1111/j.1559-1816.2004.tb02543.x

Jackson, L., Chapman, P., & Crundall, D. (2009). What happens next? Predicting other road users' behaviour as a function of driving experience and processing time. *Ergonomics, 52*(2), 154–164. https://doi.org/10.1080/00140130802030714

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*, 241–289.

Mazza, S., Gerbier, E., Gustin, M.-P., Kasikci, Z., Koenig, O., Toppino, T. C., et al. (2016). Relearn faster and retain longer along with practice, sleep makes perfect. *Psychological Science, 27*(10), 1321–1330. https://doi.org/10.1177/0956797616659930

McKenna, F. P., & Horswill, M. S. (1999). Hazard perception and its relevance for driver licensing. *Journal of the International Association of Traffic and Safety Sciences, 23*(1), 26–41.

McKenna, F. P., Horswill, M. S., & Alexander, J. L. (2006). Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology: Applied, 12*(1), 1–10. https://doi.org/10.1037/1076-898X.12.1.1

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*(2), 187–193. https://doi.org/10.3758/BF03194050

Regan, M. A., Triggs, T. J., & Godley, S. T. (2000). *Simulator-based evaluation of the DriveSmart novice driver CD-ROM training product.* Road Safety: Research, Policing & Education Conference (Handbook and Proceedings).

Reyes, M. L., O'Neal, E. (2020). *A Simulator-Based Evaluation of Two Hazard Anticipation Training Programs for Novice Drivers (Technical report).* AAA Foundation for Traffic Safety, Washington DC. Retrieved from: https://aaafoundation.org/a-simulator-based-evaluation-of-two-hazard-anticipation-training-programs-for-novice-drivers/.

Rosenbloom, T., Perlman, A., & Pereg, A. (2011). Hazard perception of motorcyclists and car drivers. *Accident Analysis and Prevention, 43*(3), 601–604. https://doi.org/10.1016/j.aap.2010.08.005

Šeibokaitė, L., Endriulaitienė, A., Markšaitytė, R., & Slavinskienė, J. (2022). Improvement of hazard prediction accuracy after training: Moderation effect of driving self-efficacy and road safety attitudes of learner drivers. *Safety Science, 151*, Article 105742. https://doi.org/10.1016/j.ssci.2022.105742

Shimazaki, K., & Ishida, T. (2009). Time series analysis of hazard detection and risk perception of accident repeaters. *Japanese Journal of Applied Psychology, 34*, 1–9.

Shinar, D. (2007). *Traffic Safety and Human Behavior.* Elsevier. http://www.knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=2909.

Smith, S. S., Horswill, M. S., Chambers, B., & Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis and Prevention, 41*(4), 729–733. https://doi.org/10.1016/j.aap.2009.03.016

Sun, L., Li, S., & Chang, R. (2019). *Development and Validity of Chinese Hazard Perception Test.* Singapore: Man-Machine-Environment System Engineering.

Taylor, T. G., Masserang, K. M., Pradhan, A. K., Divekar, G., Samuel, S., Muttart, J. W., Pollatsek, A., & Fisher, D. L. (2011). *Long term effects of hazard anticipation training on novice drivers measured on the open road.* Proceedings of the International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.

Thomas, F. D., Korbelak, K. T., Divekar, G., Blomberg, R. D., Romoser, M. R., & Fisher, D. L. (2017). *Evaluation of an updated version of the risk awareness and perception training program for young drivers*. National Highway Traffic Safety Administration (Report DOT HS 812 379). https://rosap.ntl.bts.gov/view/dot/2086.

Tūskė, V., Šeibokaitė, L., Endriulaitienė, A., & Lehtonen, E. (2019). Hazard perception test development for Lithuanian drivers. *IATSS Research, 43*(2), 108–113. https://doi.org/10.1016/j.iatssr.2018.10.001

Unverricht, J., Siby, S., & Yusuke, Y. (2018). Latent hazard anticipation in young drivers: Review and meta-analysis of training studies. *Transportation Research Record, 2672*(33), 1–9. https://doi.org/10.1177/0361198118768530

Ventsislavova, P., et al. (2019). A comparison of hazard perception and hazard prediction tests across China, Spain and the UK. *Accident Analysis and Prevention, 122*, 268–286. https://doi.org/10.1016/j.aap.2018.10.010

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A study of learner and new drivers*. Department for Transport. https://webarchive.nationalarchives.gov.uk/20101007192957/http:/www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme2/cohort2/cohortiifindings.pdf.

Wetton, M. A., Hill, A., & Horswill, M. S. (2011). The development and validation of a hazard perception test for use in driver licensing. *Accident Analysis & Prevention, 43*(5), 1759–1770. https://doi.org/10.1016/j.aap.2011.04.007

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

Willingham, D. T. (2009). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom.* Wiley.

Zhang, T., Hajiseyedjavadi, F., Wang, Y., Samuel, S., Qu, X., & Fisher, D. (2018). Training interventions are only effective on careful drivers, not careless drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 693–707. https://doi.org/10.1016/j.trf.2018.07.004