Accident Analysis and Prevention 232 (2026) 108538

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Adapting hazard perception testing to the Czech driving context: a comparison of hazard perception paradigms in differentiating driver experience and collision history

Petya Ventsislavova [a,*], Matus Sucha [b], Lydia Harrison [c], Jiri Novotny [d], David Crundall [a]

[a] *Department of Psychology, School of Social Sciences, Nottingham Trent University, UK*
[b] *Department of Psychology, Faculty of Arts, Palacký University Olomouc, Czech Republic*
[c] *School of Psychology, Faculty of Science, University of Nottingham, UK*
[d] *Driving School Association of the Czech Republic, Czech Republic*

ARTICLE INFO

Keywords:
Hazard perception
Hazard prediction
Road collisions
Driving experience
Driving context

ABSTRACT

Hazard perception (HP) testing is widely regarded as an effective approach for identifying unsafe drivers and mitigating collision risk. The European Transport Safety Council has therefore recommended its adoption within national licensing systems. Evidence indicates that HP assessments must be adapted to local environments to yield comparable safety benefits. Moreover, traditional HP tests depend largely on reaction times, which may introduce post-perceptual biases because responses typically reflect the moment a hazard is judged as threatening. To address this, the present study developed two versions of the HP test tailored to the Czech driving context: a conventional reaction-time-based test and a prediction-based test in which participants were asked to anticipate how a hazardous situation would unfold. Hazardous traffic scenarios were filmed across the Czech Republic and edited into hazard perception and prediction tests. A total of 225 participants were recruited and randomly allocated to either the hazard perception or hazard prediction test. Within each test, participants were grouped according to pre-existing driving experience (novice vs. experienced) and collision history (collision-free vs. collision-involved). The results indicated that experienced drivers outperformed novices on both test versions, showing faster and more accurate hazard detection and superior predictive performance. These findings support the feasibility of incorporating a hybrid HP assessment into the Czech national licensing system. However, significant collision group differences emerged only in the perception task, with collision-involved drivers demonstrating slower hazard response times. This suggests that the key differences may arise from post-perceptual decision-making processes rather than perceptual limitations.

1. Introduction

Since 2014, road safety improvements in the EU have saved approximately 23,800 lives, resulting in an estimated societal benefit of €60 billion. However, had the EU achieved the target 6.7% annual reduction rate, 49,590 more lives could have been saved (European Transport Safety Council). Globally, traffic collisions cause about 1.35 million deaths each year, surpassing both HIV/AIDS and tuberculosis, making this the eighth leading cause of death (World Health Organization, 2018). Further, traffic collisions are now the leading cause of death for people under 29, posing a significant risk to young populations. Beyond the human costs, road crashes also impose a substantial economic burden amounting to over $100 billion annually in low- and middle-income countries (World Health Organization, 2013). In response, the UN's *Decade of Action for Road Safety* aimed to halve global road deaths between 2010 and 2020 (United Nations, 2010) and the EU's *Vision Zero* seeks almost zero fatalities by 2050 (Cinea, 2022). Nevertheless, by 2019 the global reduction in road deaths reached only 18% and some nations, including Colombia, the Netherlands, and the US, recorded increases instead (International Transport Forum, 2022).

One promising initiative to address this issue focuses on enhancing the driver licensing process, as newly licensed drivers are disproportionately represented in crash statistics (White et al., 2011). As part of this initiative, the ETSC has recommended that hazard perception training or assessment be incorporated into all European driver licensing systems (European Transport Safety Council). In the UK, such testing

---

* Corresponding author at: Nottingham Trent University, 50 Shakespeare St, Nottingham NG1 4FQ, UK.
E-mail address: petya.petrova@ntu.ac.uk (P. Ventsislavova).

https://doi.org/10.1016/j.aap.2026.108538
Received 11 December 2025; Received in revised form 6 March 2026; Accepted 2 April 2026

reduced novice drivers' crash involvement by 17.4% in the first year (Wells et al., 2008) and following this success, similar tests have been successfully adopted in Australia, New Zealand, and the Netherlands, as well as trialled in other European countries. Despite this, most European countries have yet to implement hazard perception testing.

Hazard perception refers to a driver's ability to detect and respond to potentially dangerous situations on the road (Ventsislavova et al., 2019). Drivers with stronger hazard perception skills can anticipate and respond to potential dangers more quickly and effectively, thus reducing their likelihood of being involved in collisions. The theoretical framework most closely aligned to hazard perception is the Situation Awareness Framework (Endsley, 1995; Crundall, 2016), which conceptualises hazard perception as a three-stage process: perceiving relevant elements in the environment, understanding their meaning and possible interactions, and anticipating how the situation is likely to develop (Endsley, 1995). Empirically, hazard perception is one of the few driving skills consistently linked to collision risk (Horswill et al., 2013). It is typically assessed via short video clips filmed from the driver's perspective, during which participants are instructed to click a button whenever they identify a hazard (Horswill et al., 2004). Faster responses within the defined "hazard window" indicate better hazard perception skill as research indicates that slower response times are predictive of higher crash frequency (Horswill et al., 2004; Horswill et al., 2015). Poorer test performance has also been linked to a higher frequency of heavy-braking events in real-world driving (Hill et al., 2019) and individuals who fail hazard perception tests multiple times are nearly twice as likely to be involved in a collision after obtaining their driving licence (Boufous et al., 2011). Overall, these findings indicate that hazard perception is a strong predictor of crash risk, underscoring its importance as a critical focus for interventions aimed at enhancing driver safety and reducing collision rates.

While these results highlight the predictive validity of hazard perception measures, questions remain regarding their ecological validity. Measures obtained from HP tasks have been shown to correspond closely with comparable measures collected during real-world driving. Indeed, drivers cannot interact with video-based stimuli in the same manner as they can in driving simulators. However, it is equally the case that certain hazardous driving scenarios are difficult to recreate safely and realistically within simulator environments. Video-based HP tests therefore offer a pragmatic and ecologically relevant alternative, typically employing non-staged hazards captured in naturalistic traffic conditions. As such, they provide one of the closest available approximations to real-world hazardous situations, while remaining cost- and time-efficient. A related question concerns why relatively simple responses are used to assess driving performance in HP tests. There are several reasons for this approach. First, the validity of HP tests has been well established (e.g. (Crundall et al., 2012; Horswill and McKenna, 1999; Wetton et al., 2011), with robust evidence demonstrating an association between superior hazard perception skills and reduced collision involvement (Wells et al., 2008; Boufous et al., 2011; Drummond, 2000). For example, Horswill et al. (Horswill et al., 2015) reported that provisional drivers in Queensland who anticipated traffic conflicts earlier in a licensing HP test reported fewer crashes both in the year following the test and in the years preceding it. Comparable relationships between HP test performance and crash involvement have been reported across a range of other HP paradigms (Cheng et al., 2011; Darby et al., 2009; Horswill et al., 2010; Rosenbloom et al., 2011). Further support for the validity of HP tests comes from studies examining their relationship with on-road driving performance. In these studies, driving ability was evaluated during a single journey on public roads lasting approximately 30 to 45 min, with participants following a predefined route while being assessed by trained examiners. Although the samples differed in age, including older drivers (Ross et al., 2014; Wood et al., 2013) and younger drivers (Mills et al., 1998), all reported significant associations between examiner-rated on-road performance and hazard perception test scores.

Beyond their predictive validity, video-based HP tests offer substantial practical advantages. They can discriminate between safer and less safe drivers while remaining highly cost- and time-efficient. Consequently, HP testing has been incorporated into official driver licensing systems in both the UK and Australia, providing an efficient means of assessing drivers at scale. Evidence from these implementations suggests reductions in non–low-speed public-road crashes in which drivers accepted some degree of responsibility (Wells et al., 2008), as well as associations between HP performance and subsequent fatal and serious injury outcomes one year after the test (Boufous et al., 2011). Finally, video-based HP tests strike a balance between ecological and experimental validity. They allow specific components of the driving task to be measured under controlled conditions while exposing participants to realistic, real-world hazardous traffic situations.

Hazard perception tasks are usually validated by their ability to distinguish between novice and experienced drivers (Hill et al., 2019). While several studies have yielded robust evidence supporting the role of hazard perception ability in differentiating between drivers of varying experience levels (Horswill et al., 2004; Wetton et al., 2010), the findings across the literature remain somewhat inconsistent. Research conducted in Norway (Sagberg and Bjørnskau, 2006), Singapore (Yeung and Wong, 2015), Spain (Ventsislavova et al., 2019) and even the United Kingdom (Chapman and Underwood, 1998; Underwood et al., 2005) has failed to identify significant experience-based differences, suggesting that the test methodology may not provide a sufficiently sensitive or comprehensive assessment of hazard perception ability. One contributing factor is the lack of consistency in defining what constitutes an "experienced driver" in the literature. While the evidence indicates that driving performance tends to improve progressively with age, reaching its peak at approximately fifty-five years (Horswill and Hill, 2021), the operational definitions of an 'experienced driver' differ considerably across studies; some defining it as between three and five years of driving experience, while others define it as ten years or more (e.g., (Horswill et al., 2004; McCartt et al., 2009). This inconsistency makes cross-study comparisons challenging.

A further concern is that the contents of the various HP tests reported in academic papers differ considerably, with some HP tests better reflecting HP than others. For instance, in Yeung & Wong's study (Yeung and Wong, 2015), the video clips used did not include cues that might allow a driver to predict the development of a hazard. This likely contributed to their finding that experienced drivers did not outperform novices, as the experienced drivers did not have access to the visual cues that they would normally use to detect hazards. Similarly, stimulus design also plays a crucial role. Studies using static images, such as the Lithuanian test (Tūskė et al., 2019), found inconsistent results compared with dynamic video-based tests (Endriulaitienė et al., 2022). Static images remove the motion cues that are essential for developing situational awareness. Evidence from eye-tracking studies indicates that experienced drivers exhibit a broader distribution of attention (Underwood et al., 2003; Konstantopoulos et al., 2010) and direct their gaze towards locations where hazards are most likely to occur (Borowsky et al., 2010). These differences in behaviour are unlikely to be accurately captured using static stimuli. Furthermore, including rich video stimuli allows for the possibility that novice drivers may incorrectly react to irrelevant cues or miss early precursors of a hazard, mistakes that reflect real world driving behaviour. This additionally highlights the need to carefully define "hazard windows" within which responses will score points, as variations in how researchers define these temporal windows can obscure genuine differences in performance (Wetton et al., 2010; Ventsislavova and Crundall, 2018).

While inconsistencies in validation and methodology do not undermine the value of traditional hazard perception tests, they highlight the need for standardised guidelines. Such tests should encompass all critical components of hazard perception (detection, comprehension, and prediction), employ well-designed dynamic stimuli that incorporate precursors to potential hazards (to assess both drivers' comprehension and anticipation of developing situations) and reflect clearly specified evaluation criteria to ensure both reliability and comparability across studies.

While the traditional hazard perception test has been successfully applied in the UK (Wells et al., 2008), its suitability for use in other countries remains uncertain. Driving culture varies widely due to differences in legislation, enforcement, infrastructure, and attitudes toward risky behaviours such as speeding or drink driving (World Health Organization, 2018). Cross-cultural research indicates considerable variation in hazard perception. For example, drivers from Spain tend to perceive greater risk when evaluating traffic scenes, irrespective of their origin, compared with drivers from the United States (Sivak et al., 1988). Chinese drivers rated hazards from Spain, the United Kingdom, and China as less dangerous than UK drivers did (Ventsislavova et al., 2019). Similarly, many drivers tend to rate familiar hazards as less risky than unfamiliar ones (Bazilinskyy et al., 2020), suggesting that both culture and experience shape hazard interpretation. Evidence regarding the international validity of the test is mixed. In Brazil, experienced drivers demonstrated faster reaction times than novices (Caparelli-Daquer et al., 2017), suggesting that similar effects may be observed in other driving contexts. However, this study employed static Canadian stimuli and compared professional with non-professional drivers; limitations that may confound the findings. By contrast, research conducted in Norway (Sagberg and Bjørnskau, 2006) and Singapore (Yeung and Wong, 2015) did not distinguish between novice and experienced drivers. Cross-cultural comparisons in hazard perception showed that Malaysian drivers exhibited significantly slower response times than their UK counterparts despite similar visual fixation patterns (Lim et al., 2013). Similarly, UK drivers outperformed Chinese and Spanish drivers in both speed and accuracy (Ventsislavova et al., 2019). Notably, these differences appeared to reflect variations in hazard criteria rather than fundamental perceptual ability. Ventsislavova et al. (Ventsislavova et al., 2019) demonstrated this by presenting Chinese, UK, and Spanish drivers with identical hazardous clips across two types of hazard tests. The first test measured reaction times in the same manner as the traditional hazard perception test, while the second required drivers to predict hazards independently of whether they perceived them as hazardous. While, both tests employed identical hazard scenarios, the previously observed cultural differences were no longer evident with the hazard prediction test. These findings suggested that even in high-risk driving environments, drivers who are very familiar with that context may adopt a higher threshold than less familiar drivers for what constitutes a hazardous situation, leading to fewer or slower responses in a standard hazard perception test. Consequently, the standard hazard perception test may be influenced by cultural differences in driving norms and interpretations of risk. Effective international adaptation, therefore, requires culturally tailored stimuli and calibration to local driving conditions to ensure validity and comparability across countries.

Drivers' allocation of attention, evaluation of potential risks, and decision-making processes are shaped by top-down mechanisms informed by their prior experiences and familiarity with the driving environment (Gazzaley and Nobre, 2012). When encountering a developing hazard, a driver may exhibit either an automatic reaction or a deliberate response. A reaction represents an immediate, unplanned behaviour triggered by the sudden detection of danger while a response involves considered, goal-directed actions based upon the individual's cognitive appraisal of the situation. The latter tends to produce more adaptive outcomes, particularly when guided by effective hazard prediction and supported by anticipatory hazard management strategies (Pradhan et al., 2017). Given the rich variety of responses that drivers can apply on the road, relying solely on reaction time to measure a drivers' hazard perception offers a restricted measure. Furthermore, simple response times may introduce post-perceptual biases as a drivers' recorded response may not reflect the instant of initial recognition, but instead the moment at which the event is cognitively categorised as hazardous. This distinction is especially relevant in contexts such as developing countries with elevated collision rates, where repeated exposure to hazardous scenarios may reduce drivers' sensitivity to potential threats (e.g. (Lim et al., 2013). Consequently, it is essential to also evaluate drivers' anticipatory skills; specifically, whether their visual attention is directed to appropriate areas at the appropriate time, regardless of how they calibrate the unfolding situation.

The hazard prediction test offers a methodologically sound solution to this issue. This paradigm closely resembles the traditional hazard perception test, as it presents traffic scenarios containing hazardous situations. However, unlike the traditional method, the video clip is occluded before the hazard materialises. Drivers are then asked to predict how the situation is likely to unfold, regardless of whether they regard it as hazardous or not (e.g., (Crundall, 2016; Ventsislavova and Crundall, 2018). The hazard prediction test therefore addresses an important limitation of the traditional hazard perception test, as it requires participants to anticipate the development of a hazardous situation independently of their subjective evaluation of its risk. This isolates the predictive component of hazard perception and provides a measure of accuracy that is not confounded by subjective biases.

Early iterations of the hazard prediction test used open-response formats (Jackson et al., 2009), while more recent versions have employed multiple-choice questions (Ventsislavova and Crundall, 2018) which are both more practical to administer and equally effective at distinguishing between novice and experienced drivers. Empirical evidence indicates that this paradigm has yielded promising results. Lim et al. (Lim et al., 2013; Lim et al., 2014) reported that traditional hazard perception tests failed to capture experience-related differences among Malaysian drivers, whereas the hazard prediction test successfully differentiated between experienced and novice drivers. Ventsislavova et al. (Ventsislavova et al., 2019) observed a similar pattern across the UK, Spain, and China. Additional studies employing context-specific hazard prediction tasks have been found to be effective in Israel (Ventsislavova et al., 2022), Lithuania (Endriulaitienė et al., 2022), and Australia (Horswill et al., 2020) further corroborated its validity by distinguishing drivers according to experience and crash history.

Other researchers have instead sought to minimise criterion bias by modifying the traditional hazard perception test while still retaining its reaction-time component. For instance, Manley et al. (Manley et al., 2020) developed a Thai version that instructed participants to anticipate "traffic conflicts" rather than hazards, effectively differentiating experience levels. Similarly, Sun et al. (Sun et al., 2019) designed a Chinese hazard perception test based on national guidelines and expert validation, demonstrating that well-calibrated traditional tests can still distinguish between drivers of differing experience, even in high-risk driving cultures. Suggesting that, while the traditional hazard perception test has certain limitations when applied in countries with driving cultures that differ substantially from that of the United Kingdom, carefully designed traditional tests can still effectively distinguish between novice and experienced drivers (Manley et al., 2020; Sun et al., 2019). Consequently, researchers have recommended that both versions of the test be utilised for assessment and training purposes, and that they be piloted across different countries (e.g. (Ventsislavova et al., 2019).

The Czech Republic presents a particularly relevant context for such research. Between 2010 and 2020, the country recorded 7,372 road fatalities, the tenth highest within the European Union (European Commission, Directorate-General for Mobility and Transport, 2022). Young drivers aged 18–24 are 3.7 times more likely to be involved in fatal crashes and rank fourth highest in the EU for fatalities within this age group (European Commission, 2021). Reducing young driver casualties is identified as a key priority in the Czech Road Traffic Safety Strategy 2021–2030 (Besip, 2021). Strengthening the learning and licensing processes could play a crucial role in addressing this issue, underscoring the importance of evaluating hazard perception assessments. To date, Czech drivers have not been formally exposed to any version of the hazard perception test. Given the demonstrated effectiveness of both the traditional and prediction-based approaches in other countries, and the national commitment to improving young driver safety, the present study aims to adapt hazard perception testing to the Czech driving context and examine which version most effectively differentiates between safe and unsafe drivers. Accordingly, the study examined the capacity of the tests to discriminate between experienced and novice drivers, as well as between drivers with and without prior collision involvement, in order to evaluate whether this version of the tests similarly predicts retrospective collisions, consistent with findings from previous research (see (Boufous et al., 2011; Drummond, 2000). The overarching objective is to develop a test that can be integrated into the Czech driver licensing system as an evidence-based assessment tool.

This study aims to develop hazard prediction and hazard perception tests that are specifically adapted to the Czech driving context and to evaluate their effectiveness in distinguishing between experienced and novice drivers, as well as between collision-free and collision-involved drivers. Based on previous findings, it is anticipated that the hazard prediction test will effectively differentiate between experience groups, whereas the traditional hazard perception test may be less sensitive to such differences. However, given that only one study to date has examined the relationship between hazard prediction performance and collision involvement (Horswill et al., 2020), while hazard perception tests have consistently demonstrated the ability to distinguish between collision-involved and collision-free drivers (Wells et al., 2008; Drummond, 2000), it is expected that both tests will successfully differentiate drivers based on collision history.

2. Methods

2.1. Design

The study employed a between-subjects factorial design combining experimental and quasi-experimental methods, including hazard perception test type (reaction-time–based vs prediction-based tests), driving experience (novice vs experienced drivers), and collision history (drivers with vs without self-reported prior collisions). Hazard perception test type was experimentally manipulated through random assignment of participants to test conditions. In contrast, both driving experience and collision history represented naturally occurring participant characteristics and therefore constituted quasi-experimental grouping variables. Accordingly, comparisons involving driving experience and collision history should be interpreted as quasi-experimental, whereas comparisons between hazard perception paradigms reflected an experimental manipulation. This hybrid design allowed the study to examine how different hazard perception tests differentiate drivers as a function of experience and collision history, while maintaining experimental control over test format. Driving experience (experienced vs novice), self-reported number of collisions and test type were fitted as the independent variables. The primary dependent variables were reaction time to hazards in the hazard perception task and response accuracy in the hazard prediction task. Additional measures were collected for the hazard perception task, including accuracy of hazard detection and the number of extra hazard responses per clip. Accuracy in the hazard perception task was defined as a response occurring within a predefined temporal window surrounding the onset of the hazard. For the hazard prediction task, accuracy was determined by the participant's ability to correctly anticipate the unfolding of a hazardous situation from a set of presented response options. The study received ethical approval from the Ethics Committee of Palacký University (Czech Republic) and the College of Business, Law and Social Sciences at Nottingham Trent University (United Kingdom).

2.2. Participants

To estimate the required sample size for the study, an *a priori* power analysis was conducted using G*Power. Based on the pilot data collected for this study, an odds ratio of 0.379 was calculated. With an alpha level of 0.05 and a target power of 0.95, the analysis indicated that a minimum of 99 participants per task would be required to detect a statistically significant effect of driving experience in a multiple linear regression model.

A total of 225 participants were recruited for the study, with 119 completing the hazard perception task and 106 completing the hazard prediction task. Participants were recruited through university mailing lists, driving schools, and public advertisements targeting Czech drivers with varying levels of driving experience. Screening criteria were applied prior to participation to ensure eligibility, including possession of a valid provisional or full driving licence and active engagement in driving at the time of the study. Novice drivers comprised both newly licensed drivers and individuals currently enrolled in driving schools, reflecting early-stage drivers with limited independent driving experience. To be included in the study, experienced drivers were required to have held a full driving licence for at least three years and to have driven a minimum of 2,000 km within the preceding two years (McCartt et al., 2009). Novice drivers were required to have passed their driving test within the past year or to be currently learner drivers. While learner drivers typically accumulate relatively low mileage under supervised conditions, their inclusion was intentional, as the study aimed to capture hazard perception and prediction skills at the earliest stages of driving experience. Collision history was self-reported and referred to any traffic collision in which the participant was directly involved as a driver or learner driver, irrespective of fault. However, detailed contextual information regarding collision circumstances (e.g., supervision status, severity, or responsibility) was not collected.

Participants were recruited using a purposive sampling strategy designed to maximise variability in driving experience while maintaining ecological validity within the Czech driving context. Novice and experienced drivers were classified based on a combination of licence tenure and recent driving exposure, rather than age alone, in order to minimise misclassification bias. Although recruitment was based on voluntary participation, stringent inclusion criteria were applied to limit heterogeneity within experience groups. Participants were randomly allocated to the hazard perception and hazard prediction tasks, thereby reducing the likelihood of systematic selection effects across task conditions. Participant demographics, broken down by task type and experience group, are reported in Table 1. A *t*-test determined that there were no significant difference in the ages of participants who completed the hazard perception or hazard prediction tasks, $t(222) = 0.137$, $p = 0.616$. Similarly, there was no significant difference in the number of kilometres driven in the past two years by participants in the hazard perception and prediction tasks, $t(207) = 0.939$, $p = 0.059$.

2.3. Materials and apparatus

2.3.1. Video clips

2.3.1.1. Filming and synchronisation. To develop the video stimuli for the hazard perception and prediction tasks, approximately 17 h of bespoke driving footage were recorded across various regions in the Czech Republic in October 2022. Filming was conducted across a variety of driving environments to capture a broad spectrum of scenarios representative of those typically encountered by drivers in the Czech Republic. All footage was captured during daylight hours and under clear weather conditions, across different times of day. None of the hazardous events were staged. The footage was recorded using four GoPro HERO4 Silver Edition cameras mounted on the vehicle's interior, including the windscreen, rear-view mirror, and side mirrors to accurately replicate the visual input available during real-world driving. Recordings were made in 1920 × 1080 resolution at a frame rate of 50 frames per second (fps). The vehicle was driven by an experienced Czech driver who adhered to all relevant legal and safety requirements for filming while driving. Following data collection, footage from the four

Table 1
Demographic characteristics of all participants across task and level of driving experience.

	Hazard Perception		Hazard Prediction	
	Experienced drivers	Novice Drivers	Experienced drivers	Novice Drivers
N Participants	60	59	52	54
Average age	40.19	18.44	37.98	20.39
Age *SD*	16.32	3.04	14.11	12.80
N Gender				
Male	34	29	32	20
Female	24	30	19	32
Chose not to answer / Other	2	0	1	2
Driving Experience				
N Participants still in driving school	0	42	0	38
Average estimated months between individuals driving test and participation in the experiment* (*SD*)	263.27 (195.24)	8.25 (6.92)	209.61 (160.45)	7.55 (7.62)
Average estimated hours spent in driving school**	N/A	7.44 (8.87)	N/A	8.69 (6.52)
Average estimated km driven in the previous 2 years (*SD*)	33,476.27 (43,788.20)	1,526 (5029,5)	26,581.00 (29,086.28)	173.35 (319.58)
Collision History				
N participants who had previously experienced a collision	31	10	28	6
Average N collisions in participants who reported experience of collisions (*SD*)	1.90(1.15)	1.10 (0.30)	1.89(1.11)	1.17 (0.37)
N participants who had previously experienced a near collision	45	27	45	23
Average N near collisions in participants who reported experience of near misses (*SD*)	4.50(4.50)	3.40 (5.22)	7.31(15.49)	2.73 (2.11)

*As the novice group includes all participants who reported less than 3 years of driving experience, the calculation of participants months of driving experience after passing the driving test does not include those participants who reported still being in driving school. Further, as one experienced participant in both tasks did not report the date at which they received their driving licence, they were excluded from this calculation.

**Participants estimated hours spent in driving school was only calculated for participants who reported that they were still in driving school.

cameras was synchronised and integrated into a composite overlay representing the interior view of a vehicle, using Adobe Premiere Pro 2023 (see Fig. 1). This overlay was designed to enhance ecological validity by closely replicating the visual information available to a driver in a naturalistic driving context. From the recorded footage, an initial set of 79 hazardous scenarios was identified. A hazard was defined as an object or event that would require an evasive manoeuvre, including sudden braking or swerving to avoid a potential collision (this definition has been followed to select all of the videos in this study) (see (Ventsislavova et al., 2019). Each scenario involved a materialised hazard requiring an active driver response, such as braking or evasive manoeuvring to prevent a collision.

Fig. 1. An example frame from a hazardous clip, including the vehicle interior overlay alongside the visual information captured in the right, left, and rear-view mirrors.

2.3.1.2. Hazard perception clips. To assess participants' accuracy, a hazard window was defined for each clip. Hazard onset was defined as the moment at which sufficient visual information became available for a driver to recognise the emerging hazard and initiate an appropriate response to avoid a collision. Hazard offset was defined at the point beyond which an effective response would no longer have been possible (see (Ventsislavova et al., 2019). The duration of the pre-defined hazard windows ranged between 1,840 ms and 6,650 ms across video clips, varying by clip according to the temporal duration of each hazard. Hazard response times were calculated by subtracting the hazard onset time for each clip from the participant's button-press time. Participants who failed to respond within a given hazard window were assigned a score of 0. Responses were scored on a scale from 0 to 5, where a score of 5 indicated the fastest response within the hazard window and a score of 0 indicated no response during the hazard window. This scoring procedure followed the method established by the UK Department for Transport (DfT) for official hazard perception testing, with scores determined by the speed of hazard detection. Participants who produced an excessive number of responses in a single clip (i.e., more than seven clicks) were excluded from the analysis.

One example of a hazardous scenario involved the driver travelling in the middle lane of a motorway, while a truck in the adjacent right-hand lane flashed its indicator and began merging into the same lane. This situation required the driver to either brake or change lanes to the left in order to avoid a potential collision (see Fig. 2). Participants were provided with a definition of a hazardous situation (*see Section 2.3.1.1: Filming and Synchronisation*) and were instructed to observe each video clip carefully, pressing the mouse button whenever they believed they had identified a hazard.

2.3.1.3. Hazard prediction clips. Each hazard prediction clip contained a hazard that would subsequently materialise, but each clip was occluded

Fig. 2. The image depicts a motorway from a driver's perspective. A car and a white truck are visible ahead. The truck is partially encroaching from the right-hand lane into the driver's lane. The merging manoeuvre by the truck requires the driver to either reduce speed or change lanes to avoid a potential collision.

by a black screen just as the hazard began to develop. The occlusion point was carefully selected so that a relevant cue indicating the forthcoming hazard was visible, allowing participants to anticipate how the situation was likely to unfold. For example, in the scenario shown in Fig. 2, the clip was occluded shortly after the second intermittent indicator signal, providing a clear precursor to the hazard for participants to anticipate the truck's merging manoeuvre. To prevent participants from using the final frame of the clip to search for the hazard at their leisure, each clip included a one-second black screen following the occlusion. This occlusion technique has been previously identified as the best way to elicit experienced-novice driver differences (Jackson et al., 2009). After viewing each clip, participants were asked to select the correct answer out of four multiple-choice response options. One option accurately described the development of the hazard following the occlusion, while the remaining three served as distractors. While all distractor options were contextually feasible within the given driving scenario, they were designed such that, for an experienced driver, only the correct option could be logically inferred from the preceding hazard precursor. The multiple-choice response options were carefully constructed and piloted to ensure that none appeared more plausible than the others based on linguistic features such as wording, phrasing, or sentence length (see (Ventsislavova and Crundall, 2018).

2.3.1.4. Clip selection and focus group. Following an initial screening, 37 video clips were shortlisted for further evaluation. A focus group was organised comprising nine experts in transport psychology and road safety to assess the clips. Participants were recruited from both the United Kingdom (n = 3) and the Czech Republic (n = 6), and all attended the session in person. Each of the 37 clips was presented in two formats: a hazard prediction version (in which the video occluded at the point of hazard onset), and a hazard perception version (in which the clip was shown in full, revealing the materialised hazard). During the session, participants were shown all 37 video clips, beginning with the hazard prediction version of each clip and asked whether they were able to predict the unfolding hazard. This was followed by the corresponding hazard perception version, after which participants were asked to evaluate the plausibility of the occlusion point and the adequacy of the overall hazard window. Finally, participants were asked to evaluate whether the depicted hazard was representative of typical driving situations in the Czech Republic. The discussion centred on several key criteria: whether sufficient information was available for participants to accurately predict the development of the scenario; whether the occlusion point occurred too early or too late; whether the distractor options were plausible; whether the wording of the multiple-choice options was potentially misleading or biased; and whether the overall quality of the video footage was adequate for inclusion in the tests.

For the hazard prediction task, participants responded by selecting one of four multiple-choice options using OMBEA Response clickers. This allowed the plausibility of the generated distractors to be assessed. Once all participants had submitted their responses, a summary showing the proportion of responses for each option was displayed. The correct answer was then revealed, alongside a screenshot from the final frame of the occluded clip to illustrate the visual cues that could have supported correct prediction. This procedure is illustrated in Fig. 3.

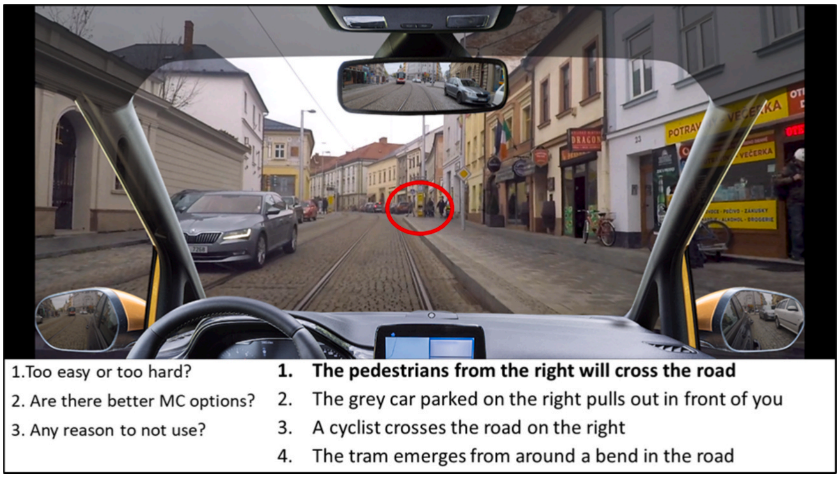
To minimise order effects, both the sequence of clip presentation and the order of multiple-choice response options were randomised. This approach ensured that similar hazard types were not grouped in a way that might influence participants' accuracy.

Following the focus group, 31 clips were selected for piloting to determine the most suitable stimuli for inclusion in the final hazard prediction and hazard perception tasks. The other six clips were excluded if our driver's behaviour was considered inconsistent with standard practice (n = 3), the scenario was not perceived as hazardous (n = 2), or the simultaneous occurrence of multiple hazards within a single clip would prevent the clip being used in a hazard perception test (n = 1). Based on the feedback from the focus group, 13 clips were re-edited to adjust the timing of the occlusion point, and two clips were shortened to reduce the time leading up to the onset of the hazard. Additionally, the multiple-choice options for 29 clips were revised to enhance plausibility and clarity, ensuring cultural and contextual appropriateness for a Czech participant base. Excluding the one-second black occlusion screen, the final 31 hazard prediction clips ranged in duration from 11.70 s to 38.74 s (M = 29.72 s), while the corresponding hazard perception clips ranged from 24.44 s to 50.68 s (M = 44.86 s). The focus group session took place over two consecutive days and lasted a total of ten hours.

2.4. Pilot study

The 31 video clips shortlisted during the focus group were subjected to a pilot study to determine which were most effective at differentiating between novice and experienced drivers. The study also aimed to evaluate the difficulty level of each clip and to ensure that none exhibited ceiling or floor effects that could compromise the sensitivity of the measures. A total of 19 participants (10 experienced drivers, 9 novice drivers) completed the hazard perception pilot, while 21 participants (12 experienced drivers, 9 novice drivers) took part in the hazard prediction pilot. Demographic characteristics of the pilot sample are presented in Table 2. All participants were native Czech speakers.

The pilot study was developed using Gorilla Experiment Builder and



1.Too easy or too hard?
2. Are there better MC options?
3. Any reason to not use?

1. The pedestrians from the right will cross the road
2. The grey car parked on the right pulls out in front of you
3. A cyclist crosses the road on the right
4. The tram emerges from around a bend in the road

Fig. 3. An example slide shown to participants following a multiple-choice question. Participants were shown the final frame of a hazard prediction clip, highlighting the cues present within the clip. The correct multiple-choice option was bolded and explained to participants.

Table 2
Demographic information of participants recruited for the pilot study across driving experience and task type.

	Hazard perception task		Hazard prediction task	
	Experienced	Novice	Experienced	Novice
N Participants	10	9	12	9
Average Age	30.10	19.22	32.83	19.11
Age *SD*	8.24	0.44	7.84	0.33
Gender				
N Female	5	7	5	7
N Male	5	2	7	2
Level of Experience				
Average number of months since test passed	131.50	9.43	139.17	8.5
SD	97.83	1.90	93.70	3.25
Average estimated Km driven in previous two years	17,900	912.50	22863.64	688.29
SD	19180.14	1387.89	22724.54	947.58
Collision involvement				
N indicated collision involvement	4	2	6	0

was administered in person. All instructions and on-screen materials were presented in Czech. The procedure for the pilot study was identical to that of the final experimental protocol for both the hazard perception and hazard prediction tasks (see *Section 2.5 Main Study Procedure* for details). The video stimuli were identical across both tasks, with the only difference being that the hazard fully materialised in the hazard perception clips. To evaluate the discriminatory power of each clip, Wilcoxon rank-sum tests were conducted to compare performance between novice and experienced drivers. These analyses were carried out separately for each clip. Further details on the derivation of reaction time and accuracy measures for each task are provided in Section 3.

2.4.1. Pilot results

The results of the pilot study indicated significant differences in mean accuracy between experienced drivers ($\mu = 0.81$) and novice drivers ($\mu = 0.62$) on the hazard prediction task, $W = 61{,}845$, $p < 0.001$, $r = 0.213$. Similarly, for the hazard perception task, the mean reaction times of experienced drivers ($\mu = 2.18$ s) were significantly faster than those of novice drivers ($\mu = 1.68$ s), $W = 50{,}012$, $p = 0.0007$, $r = 0.139$. However, no significant difference was observed between novice and experienced drivers in terms of hazard detection accuracy, $W = 45{,}152$, $p = 0.264$, $r = 0.046$.

A clip-by-clip analysis was also conducted to identify which video stimuli most effectively distinguished between novice and experienced drivers. The results of this analysis are presented in Table 3. Based on this analysis, six hazard prediction clips and three hazard perception clips demonstrated statistically significant differences in performance between the two groups. An additional six hazard prediction clips and three hazard perception clips did not reach statistical significance but yielded moderate to large effect sizes, suggesting potential utility in discriminating driver experience.

The final selection of video clips for the main experimental procedure was determined by a steering group composed of six transport psychology experts, including four based in the Czech Republic and two in the United Kingdom. Clip selection was guided by the results of the clip-by-clip analysis, differences in mean reaction times and accuracy scores between experience groups, and the contextual relevance and visual quality of the hazards depicted. This process resulted in the selection of 18 clips for use in both the hazard prediction and hazard perception tasks. While previous research has recommended the use of identical video stimuli across both hazard perception and hazard prediction tasks to ensure that any observed differences are attributable to task demands rather than variations in the stimuli (e.g., (Ventsislavova et al., 2022), this was only partially feasible in the present study. Twelve of the selected clips were identical across both tasks, while six differed. These non-overlapping clips were excluded from cross-use based on the clip-by-clip analysis, which indicated that they were not equally suitable for both task types. As noted in prior research, it is not always possible to use identical clips across both methodologies (e.g., (Crundall, 2016). Certain hazards develop too abruptly to include a meaningful precursor, rendering them unsuitable for hazard prediction, while others include clear precursors but do not yet materialise into an imminent threat, making them inappropriate for hazard perception tasks. The final evaluation indicated that only one video clip required adjustment to its hazard window. All other clips were retained in their original form, as established during the pilot phase. The complete set of stimuli selected for inclusion in the hazard perception and hazard prediction tasks is detailed in Table 4.

2.5. Main study procedure

The main experimental procedure was implemented using Gorilla Experiment Builder and conducted in person. The experiment was conducted in a controlled laboratory environment to ensure standardised testing conditions across participants. All stimuli were presented on a 24-inch LCD monitor with a native resolution of 1920 × 1080 pixels and a refresh rate of 60 Hz. Participants were seated at an approximate viewing distance of 60–70 cm from the screen, resulting in a visual angle comparable to that of a natural driving scene. The laboratory was evenly illuminated with ambient lighting to minimise screen glare and visual fatigue, and external noise was minimised throughout testing. Participants completed the task individually, using a standard computer mouse for responses. All experimental sessions followed identical hardware and environmental configurations to ensure consistency in response measurement across participants. The same equipment and setup were used for both the hazard perception and hazard prediction tasks.

Upon arrival, participants were provided with detailed information about the study and asked to give informed consent. Following consent, participants were asked to complete a brief questionnaire capturing their demographic information and driving experience. Participants were then randomly assigned to either the hazard prediction or the hazard perception task. Prior to the experiment, participants completed a practice trial to familiarise themselves with the procedure and were given the opportunity to ask any questions. The practice trial depicted a hazardous situation that was only presented during the practice and did not include any of the clips or hazards used in the main test. The entire study, including all instructions and stimuli, was presented in Czech, which was the participants' native language. None of the participants had previously completed any form of hazard perception or prediction testing.

2.5.1. The hazard prediction test

The independent variables consisted of the drivers' experience levels and their history of involvement in collisions, while the dependent variable was the accuracy of hazard prediction. Participants were informed that they would be shown 18 video clips depicting real driving scenarios filmed from the driver’s perspective on roads in the Czech Republic. They were instructed that each clip would be occluded by a black screen immediately prior to a hazardous scenario. Following each clip, participants were asked what would happen next in the driving scene and asked to select the correct answer from four multiple-choice options. Participants were informed that only one of the options accurately described the actual development of the hazardous situation depicted in the video, and that there would always be sufficient visual information in the environment prior to occlusion to allow an attentive observer to correctly predict the outcome. Before the experimental trials began, participants completed a practice trial designed to familiarise them with the procedure. Performance feedback was provided only during the practice trial, during which the clip was replayed with the hazard highlighted to assist participants in identifying the relevant cues

Table 3
Results of the Wilcoxon rank-sum tests for each individual clip across hazard task. The table includes, mean differences, p-values, and effect size for each.

	Hazard Prediction				Hazard Perception			
Clip Name	Difference between experienced and novice drivers' *M* accuracy rates	Wilcoxon rank-sum test results	Significant Difference?	Included in final procedure?	Difference between experienced and novice drivers' *M* reaction times	Wilcoxon rank-sum test results	Significant Difference?	Included in final procedure?
Clip1	0.028	$W = 55.5, p = 0.889, r = 0.0457$	No	No	1.344	$W = 59.5, p = 0.185, r = 0.315$*	No	No
Clip2	0.444	$W = 78, p = 0.0142, r = 0.547$	Yes – Large effect size	**Yes**	1.011	$W = 59, p = 0.248, r = 0.275$	No	**Yes**
Clip3	0.111	$W = 60, p = 0.29, r = 0.252$	No	No	0.056	$W = 48.5, p = 0.801, r = 0.0675$	No	No
Clip4	0.333	$W = 72, p = 0.0404, r = 0.460$	Yes – moderate effect size	**Yes**	0.156	$W = 49, p = 0.76, r = 0.0801$	No	No
Clip5	0.361	$W = 73.5, p = 0.0675, r = 0.409$*	No*	**Yes**	0.322	$W = 50.5, p = 0.656, r = 0.112$	No	**Yes**
Clip6	0.583	$W = 85.5, p = 0.00704, r = 0.598$	Yes – large effect size	**Yes**	0.333	$W = 45, p = 1, r = 0$	No	No
Clip7	0.111	$W = 60, p = 0.29, r = 0.252$	No	**Yes**	0.667	$W = 54, p = 0.468, r = 0.176$	No	**Yes**
Clip8	0.611	$W = 87, p = 0.00711, r = 0.596$	Yes – large effect size	**Yes**	0.767	$W = 63.5, p = 0.0992, r = 0.389$*	No	**Yes**
Clip9	−0.056	$W = 51, p = 0.77, r = 0.0767$	No	No	0.278	$W = 51.5, p = 0.615, r = 0.125$	No	**Yes**
Clip10	0.139	$W = 61.5, p = 0.566, r = 0.134$	No	**Yes**	1.233	$W = 65.6, p = 0.0828, r = 0.408$*	No	**Yes**
Clip11	0.222	$W = 66, p = 0.108, r = 0.366$*	No	**Yes**	1.144	$W = 70, p = 0.0395, r = 0.482$	Yes – moderate effect size	**Yes**
Clip12	−0.056	$W = 51, p = 0.77, r = 0.0767$	No	No	0.833	$W = 59.5, p = 0.242, r = 0.278$	No	**Yes**
Clip13	0.000	$W = 54, p = 1, r = 0$	No	No	0.656	$W = 57.5, p = 0.256, r = 0.271$	No	**Yes**
Clip14	0.250	$W = 67.5, p = 0.175, r = 0.307$*	No	**Yes**	−1.089	$W = 35, p = 0.409, r = 0.199$	No	No
Clip15	−0.028	$W = 52.5, p = 0.923, r = 0.0315$	No	No	−0.389	$W = 39, p = 0.637, r = 0.118$	No	No
Clip16	0.111	$W = 60, p = 0.643, r = 0.110$	No	**Yes**	0.333	$W = 48, p = 0.828, r = 0.0599$	No	**Yes**
Clip17	−0.083	$W = 49.5, p = 0.736, r = 0.0829$	No	No	0.411	$W = 50.5, p = 0.677, r = 0.105$	No	No
Clip18	0.361	$W = 73.5, p = 0.0675, r = 0.409$*	No*	**Yes**	0.022	$W = 54.5, p = 0.447, r = 0.184$	No	No
Clip19	0.528	$W = 82.5, p = 0.0216, r = 0.510$	Yes – large effect size	**Yes**	1.622	$W = 73, p = 0.0223, r = 0.534$	Yes – large effect size	**Yes**
Clip20	0.222	$W = 66, p = 0.341, r = 0.217$	No	**Yes**	0.722	$W = 71, p = 0.028, r = 0.514$	Yes − large effect size	**Yes**
Clip21	0.250	$W = 67.5, p = 0.286, r = 0.242$	No	**Yes**	0.522	$W = 55, p = 0.386, r = 0.209$	No	**Yes**
Clip22	0.028	$W = 55.5, p = 0.889, r = 0.0457$	No	No	−0.067	$W = 45.5, p = 1, r = 0.00952$	No	No
Clip23	−0.028	$W = 52.2, p = 0.934, r = 0.0271$	No	No	0.300	$W = 55, p = 0.398, r = 0.204$	No	**Yes**

(*continued on next page*)

Table 3 (*continued*)

	Hazard Prediction				Hazard Perception			
Clip Name	Difference between experienced and novice drivers' M accuracy rates	Wilcoxon rank-sum test results	Significant Difference?	Included in final procedure?	Difference between experienced and novice drivers' M reaction times	Wilcoxon rank-sum test results	Significant Difference?	Included in final procedure?
Clip24	0.333	$W = 72$, $p = 0.151$, $r = 0.322$*	No	**Yes**	0.411	$W = 54$, $p = 0.459$, $r = 0.18$	No	No
Clip25	0.361	$W = 73.5$, $p = 0.0675$, $r = 0.409$*	No*	**Yes**	0.511	$W = 56.5$, $p = 0.337$, $r = 0.230$	No	**Yes**
Clip26	−0.056	$W = 51$, $p = 0.77$, $r = 0.0767$	No	No	0.711	$W = 56$, $p = 0.381$, $r = 0.211$	No	**Yes**
Clip27	0.556	$W = 84$, $p = 0.00453$, $r = 0.630$	Yes – large effect size	**Yes**	0.567	$W = 54$, $p = 0.474$, $r = 0.174$	No	No
Clip28	0.028	$W = 55.5$, $p = 0.889$, $r = 0.0457$	No	No	0.767	$W = 55$, $p = 0.398$, $r = 0.204$	No	**Yes**
Clip29	0.194	$W = 64.5$, $p = 0.385$, $r = 0.199$	No	**Yes**	0.744	$W = 54$, $p = 0.4$, $r = 0.204$	No	**Yes**
Clip30	−0.056	$W = 51$, $p = 0.77$, $r = 0.0767$	No	No	−0.078	$W = 43$, $p = 0.9$, $r = 0.0386$	No	No
Clip31	0.139	$W = 61.5$, $p = 0.412$, $r = 0.192$	No	No	0.667	$W = 61$, $p = 0.125$, $r = 0.364$*	No	No

*Clips that did not reach statistical significance but yielded moderate to large effect sizes.

they may have missed. No feedback was provided during the main experimental trials. Each video began with a fixation screen lasting 3000 ms, which included a countdown to indicate when the clip would start. The order of video presentation was randomised and the clips were divided into two blocks of nine, with a short break offered between blocks. To control for potential order effects, the presentation order of the four multiple-choice response options following each clip was also randomised. The full procedure, including instructions and practice, lasted approximately 40 min. Upon completion of the task, participants were presented with a debrief screen thanking them for their participation. A visual overview of the experimental procedure is presented in Fig. 4.

2.5.2. The hazard perception test

The independent variables were driving experience and involvement in collisions. The dependent variable was reaction time to hazards. Participants were presented with 18 video clips depicting driving scenarios filmed from the driver's perspective on roads in the Czech Republic. They were informed that these clips would contain potentially hazardous situations requiring driver intervention to avoid a collision. Participants were instructed to view each clip as if they were the driver and to press the left mouse button every time they think they have seen a hazard that, in their judgment, requires an evasive action. Before beginning the experimental trials, participants completed a practice trial to ensure they understood the task requirements. Participants received feedback only during the practice trial where the clip was replayed with the hazard visually highlighted. Following the practice trial, participants proceeded to the main task without further feedback. Each trial began with a 3000 ms fixation screen, which included a countdown indicating the time remaining before the video clip commenced. All 18 clips were subsequently presented in a randomised order, divided into two blocks of nine, with an opportunity to take a short break between blocks. The duration of the full experimental was approximately 35 min. A visual representation of the procedure is provided in Fig. 4.

3. Results

3.1. Statistical analysis

The project aimed to assess the differences between experienced and novice drivers in their skills to predict and perceive hazards while driving. Data on car collisions was also incorporated into the analysis, however, the results should be interpreted with caution, as the collision groups were not equivalent in numbers. Traditionally, factorial designs of this nature have been analysed using mixed ANOVA, but this approach has several limitations. First, it treats stimuli (video clips) as a fixed factor rather than a random factor, even though the design includes two fully crossed random factors—participants and clips. Failing to account for the second random factor may increase the likelihood of Type I errors. In contrast, treating both clips and participants as random effects enhances the generalisability of the findings beyond the specific clips used in the study. Secondly, mixed ANOVA treats discrete outcomes, such as correctly identifying a hazard, as continuous variables. A more suitable approach is a multilevel generalised linear model, which allows for discrete responses and incorporates fully-crossed random factors (Baguley, 2012). In this study, we used the free and open-source software R (R Core Team, 2018). We fitted a series of models, beginning with an intercept-only model including random effects, followed by models incorporating all main effects and subsequently adding higher-order interactions, starting with two-way interactions. Likelihood ratio tests were employed to assess the significance of effects by comparing models with and without the effect of interest at each level of interaction.

Prior to the main analyses, the internal consistency of both tests was assessed and found to be acceptable. For the hazard perception test, internal consistency was high for both accuracy scores (Cronbach's $\alpha = 0.86$, 95% CI [.82, 0.90]) and reaction time measures (Cronbach's $\alpha = 0.86$, 95% CI [.82, 0.90]). The hazard prediction test also demonstrated good internal consistency (Cronbach's $\alpha = 0.82$, 95% CI [.76, 0.86]).

3.2. Hazard perception test results

There were three main measures of interest: the response times to

Table 4
Final selection of video clips for the hazard perception and hazard prediction tasks.

Hazard Perception		**Hazard Prediction**	
Clip	**Description**	Clip	**Description**
Clip2	You are driving down the middle lane of a motorway, behind a black car. A truck ahead in the right-hand lane begins to merge left into the middle lane, forcing you to pull into the left-hand lane	Clip2	You are driving down the middle lane of the motorway, behind a black car. A truck ahead in the right-hand lane begins to merge left into the middle lane, forcing you to pull into the left-hand lane. The clip occludes after the truck flashes its second indicator
Clip5	You are driving on the right-hand side of a relatively busy motorway, before pulling into the left-hand lane. A truck begins to signal and pulls into the left-hand lane, in front of you forcing you to slow down	Clip5	You are driving on the right-hand side of a relatively busy motorway, before pulling into the left-hand lane. A truck begins to signal and pulls into the left-hand lane, in front of you. The clip occludes after the truck flashes its first indicator
Clip7	You are driving down a wide city street, approaching a four-way junction. A car pulls across the junction ahead of you, causing you to brake and slow down	Clip7	You are driving down a wide city street, approaching a four-way junction. A car pulls across the junction ahead of you, causing you to brake and slow down. The clip occludes as the car becomes visible approaching the road
Clip8	You are driving down a dual carriageway, when suddenly a pedestrian appears and crosses the road ahead of you forcing you to brake and slow down	Clip8	You are driving down a dual carriageway, when suddenly a pedestrian appears and crosses the road ahead of you forcing you to brake and slow down. The clip occludes just as the pedestrian becomes visible, emerging from behind the oncoming cars
Clip10	You are driving down a residential road with cars parked on both sides. A car unexpectedly turns into the road, causing you to brake	Clip10	You are driving down a residential road with cars parked on both sides. A car unexpectedly turns into the road, causing you to brake. The clip occludes just as the car comes into view from around the corner
Clip11	You are driving in a shared space for pedestrians, trams and vehicles. A pair of pedestrians suddenly cross the road ahead of you causing you to brake and slow down	Clip11	You are driving in a shared space for pedestrians, trams and vehicles. A pair of pedestrians suddenly cross the road ahead of you causing you to brake and slow down. The clip occludes as the pedestrians are just about to step off the pavement
Clip16	You are following a grey car along a residential road. You are driving around a parked van when a cyclist that had previously been obscured by the car in front becomes visible and you must brake.	Clip16	You are following a grey car along a residential road. You are driving around a parked van when a cyclist that had previously been obscured by the car in front becomes visible and you must brake. The clip occludes just as the car moves
Clip19	You are driving behind a red car on a city street, turning left and then right. Ahead, two pedestrians on opposite sides of the road cross a zebra crossing and you must stop	Clip19	You are driving behind a red car on a city street, turning left and then right. Ahead, two pedestrians on opposite sides of the road cross a zebra crossing and you must stop. The clips occludes as the pedestrians are just about to step off the pavement
Clip20	You are driving in a shared space for pedestrians, trams and vehicles. A pedestrian suddenly emerges from a side road without looking and crosses the road forcing you to slow down	Clip20	You are driving in a shared space for pedestrians, trams and vehicles. A pedestrian suddenly emerges from a side road without looking and crosses the road forcing you to slow down. The clips occludes as the pedestrian approaches the road
Clip21	You are driving along a wide city street and approaching a crossroads where a tram is passing across. As you prepare to turn left, a car previously hidden by the tram suddenly appears and crosses your path, forcing you to brake.	Clip21	You are driving along a wide city street and approaching a crossroads where a tram is passing across. As you prepare to turn left, a car previously hidden by the tram suddenly appears and crosses your path, forcing you to brake. The clip occludes as the car briefly comes into view, emerging from behind the passing tram
Clip25	You are traveling through a residential area behind a white car. After taking the second exit at a roundabout, a bus at a nearby bus stop pulls out in front of you, forcing you to slow down.	Clip25	You are traveling through a residential area behind a white car. After taking the second exit at a roundabout, a bus at a nearby bus stop pulls out in front of you, forcing you to slow down. The clip occludes after the bus flashes its first indicator
Clip29	You are driving along cobblestone streets and turn into a side street. A parked vehicle on the corner obstructs your view, and an oncoming car suddenly becomes visible, prompting you to brake.	Clip29	You are driving along cobblestone streets and turn into a narrow side street. A parked vehicle on the corner obstructs your view, and an oncoming car suddenly becomes visible, prompting you to brake. The clip occludes just as the headlights of the oncoming car become visible from behind the parked vehicle.
Clip9	You are driving in a shared space for pedestrians, trams and vehicles. A car quickly pulls in front of you from a side road on the right forcing you to slow down	Clip4	You are driving down the left-hand lane on the motorway directly behind a white car and a truck. A grey car in the right-hand lane suddenly pulls in front of you, undertaking and cutting across your path. The clip occludes just as the grey car becomes visible through your right window and activates its indicator.
Clip12	You pull into a side road. A parked car is pulling out in the side road which only becomes visible once you have started turning causing you to brake.	Clip6	You are driving down a city road passing a tram stop. A pedestrian suddenly emerges from the right-hand side and starts crossing the road ahead. The clip occludes as the pedestrian steps off the pavement.
Clip13	You exit a side road and turn into a main road. A pair of pedestrians with a service dog cross the road ahead of you, causing you to slow down	Clip14	You are travelling behind a white car on a quiet single carriageway. A cyclist on the right-hand side of the road causes the car ahead of you to brake and pull into the left-hand lane to pass it and you to slow down. The clip occludes with the car in front braking abruptly
Clip23	You are traveling through a residential area. Firemen using a ladder are visible ahead on the road opposite a	Clip18	You are driving through a quiet, wide city street and enter a car park. A pedestrian emerges from the left and

(continued on next page)

Table 4 (*continued*)

Hazard Perception		Hazard Prediction	
	corner in the road. As you turn the corner a parked fire truck becomes visible forcing you to slow down.		walks in front of you forcing you to brake. The clip occludes as the pedestrian becomes visible from behind the cars.
Clip26	You are driving through a wide city street behind a line of traffic. As you approach a zebra crossing, two pedestrians step out and begin to cross, prompting you to stop. Just after they pass, an e-scooter suddenly appears from the side and crosses the road unexpectedly, forcing you to remain stopped.	Clip24	You are travelling through a residential area behind a line of traffic. The car in front of you begins to move into the left-hand lane, revealing a stationary car pulled over on the right, which forces you to slow down. The clip occludes just as part of the stationary car's hazard lights becomes visible
Clip28	You are driving down a narrow road and turn left into a wider shared space for pedestrians, trams and vehicles A man is seen on the right about to step out of a parked car, forcing you to reduce your speed	Clip27	You are driving down a city road and turn left into a narrow side street that leads onto another main road. A large bus coming from the left suddenly becomes visible, attempting to turn into the narrow street, forcing you to stop. The clip occludes just as a flash of the bus becomes visible.

hazards (Reaction Times – RTs), the proportion of hazards correctly identified (hazard perception accuracy), and the number of mouse clicks outside the hazard window.

3.2.1. *Differences between driving experience and collision groups in relation to their reaction times to the pre-defined hazards*

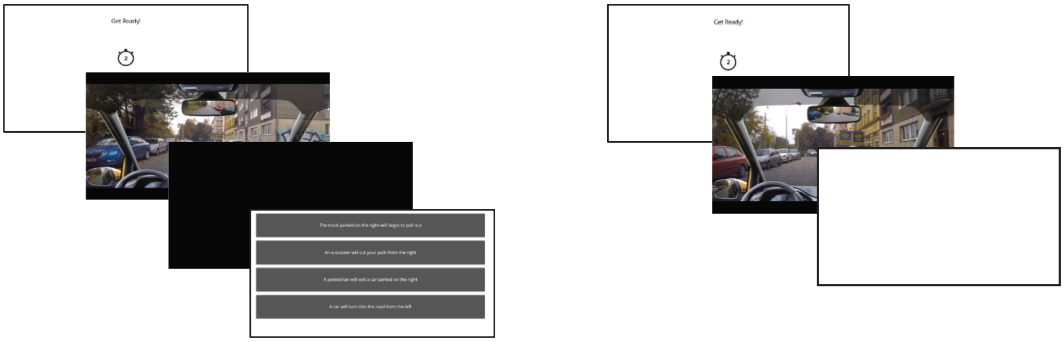
Differences for reaction times (RTs) were analysed via a multilevel ordinal regression using the R ordinal package (Christensen, 2018). In order to calculate response times for the hazards, hazard onsets and offsets were defined for each clip. RTs were then calculated by subtracting the hazard onset times for each video clip from the participants' button-press times. If participants did not respond during a particular clip, they were assigned a score of 0. The scoring system ranged from 0 to 5, with 5 indicating the fastest response within the hazard window and 0 representing no response within the hazard window (following the procedure established by the UK Department for Transport (DfT) for official hazard perception testing). The scores were determined based on the speed at which participants detected the hazard. For instance, a score of 1 would indicate a response made just before the end of the hazard window (when the hazard window closes), while a score of 5 would signify a response made at the beginning of the hazard window (when the hazard window opens). The independent variables were driving experience and involvement in collisions and the dependent variable was reaction times to hazards.

First an intercept-only model with random intercepts for participants (SD = 1.33) and video clips (SD = 0.99) was fitted before adding main effects and two-way interactions. The intercept only model (χ^2 (7) = 6388.2) was a worse fit than a model with all main effects ($\Delta\chi^2$ (2) = 27.32, $p < 0.001$), which however was a better fit than a model with all two-way interactions ($\Delta\chi^2$ (1) = 0.74, p = 0.38). The results suggested that experienced drivers were significantly faster at detecting hazards than the novices (M = 3.01 vs. M = 2.1) ($\Delta\chi^2$ (1) = 27.24, $p < 0.001$) (see Fig. 5). The results also indicated that those drivers who have never been involved in collisions were slightly faster than those who have previously been involved in collisions (M = 2.70 vs. M = 2.23) ($\Delta\chi^2$ (1) = 5.27, $p < 0.05$) (see Fig. 6). Collision-involved participants showed lower odds of achieving higher RTScore categories compared to collision-free participants (OR = 0.55, 95% CI [0.33, 0.90]), while novice drivers showed substantially lower odds than experienced drivers (OR = 0.26, 95% CI [0.16, 0.42]). These effects indicate reduced hazard perception performance associated with collision history and less driving experience. It is important to note that the collision group consisted of only 41 participants, compared to 78 in the non-collision group (combining experienced and novice drivers in the hazard perception task), warranting caution when interpreting the findings. Within the experienced driver group specifically, those who had not been involved in collisions demonstrated faster RTs (non-collision: M = 3.21 vs. collision: M = 2.83). No significant interaction was found between collision involvement and driving experience (see Fig. 6).

3.2.2. *Differences between driving experience and collision groups in relation to their hazard perception accuracy*

To assess overall hazard perception accuracy, participants' responses within the hazard window were analysed, regardless of how quickly they responded. A score of 1 was assigned if the participant detected the hazard within the pre-defined hazard window, while a score of 0 indicated a failure to respond during that time frame. The independent variables were driving experience and involvement in collisions and the dependent variable was accuracy in detecting hazards.

A 2x2 factorial model was fitted via a multilevel logistic regression using the R lme4l package (Bates et al., 2015). The deviance (likelihood ratio Chi Square, χ^2) for the intercept only model was 2466.0 and decreased dramatically to 2440.7 for a model including main effects of



Panel A **Panel B**

Fig. 4. A and B: Sequence of events in the hazard prediction task (Panel A) and hazard perception task (Panel B). Panel A illustrates the procedure for the hazard prediction task, showing the fixation screen, the final visible frame of the video clip prior to occlusion, a one-second black occlusion screen, and the subsequent display of four multiple-choice response options. The response screen remained visible until a selection was made. Panel B depicts the sequence for the hazard perception task, including the fixation screen, a representative frame from the clip showing the developing hazard, and the blank screen presented for 500 ms following the clip.

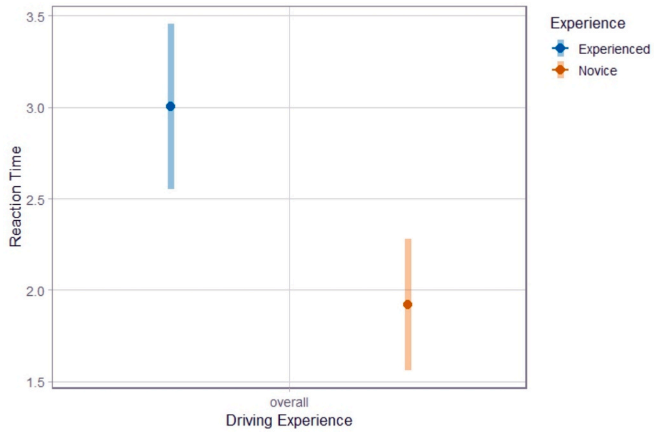


Fig. 5. Reaction Times to hazards across driving experience groups. Scores closer to 5 indicate a faster reaction time (with error bars).

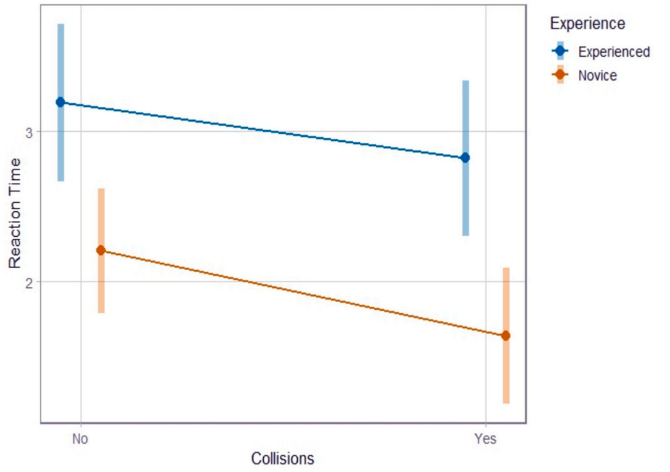


Fig. 6. Reaction time to hazards across experience and collisions groups. Scores closer to 5 indicate a faster reaction time (with error bars).

collisions and driving experience. This improvement in model fit was statistically significant, $\Delta\chi^2$ (2) = 25.37, p < 0.001. In addition, although the chi-square value decreased when all two-way interactions were included, the change was not significant, $\Delta\chi^2(1) = 1.03$, p = 0.31. The main effects model appeared to be the most informative. There was a significant difference between the experienced groups in hazard perception accuracy $\Delta\chi^2$ (1) = 25.33, p < 0001, with experienced drivers showing better HP accuracy than the novices ($M = 75\%$ vs $M = 42.3\%$) (see Fig. 7). On average, those drivers that were involved in collisions were less accurate in spotting hazards than those that have never been involved in collisions $\Delta\chi^2$ (1) = 4.58, p < 0.05, ($M = 67\%$ vs $M = 52\%$) (see Fig. 8). The odds of a correct hazard perception response were significantly lower for participants reporting a prior collision (OR = 0.54, 95% CI [0.31, 0.94]), indicating that collision-involved drivers were less likely to respond correctly than collision-free drivers. In addition, novice drivers exhibited substantially lower odds of correct hazard perception compared with experienced drivers (OR = 0.24, 95% CI [0.14, 0.41]). The model intercept corresponded to the odds of correct responding for experienced, collision-free drivers (OR = 4.21, 95% CI [2.22, 7.99]). No significant interactions between any of the groups was observed. It is important to note that the collision group consisted of only 41 participants, compared to 78 in the non-collision group, warranting caution when interpretating the results.

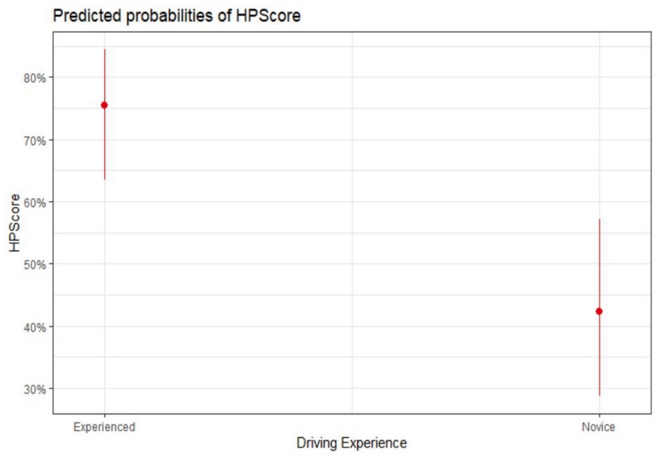


Fig. 7. Hazard perception accuracy across driving experience groups (with error bars).

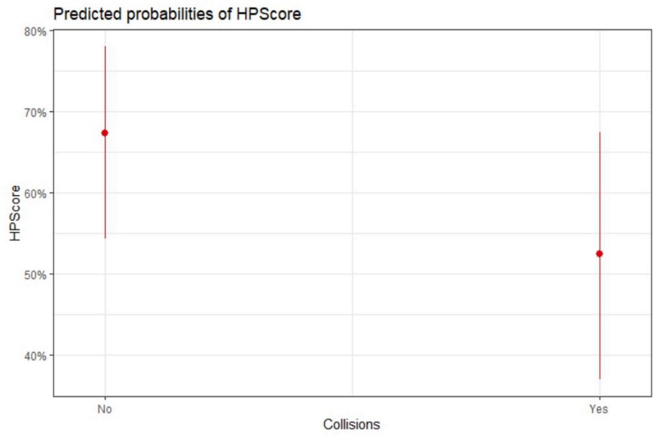


Fig. 8. Accuracy in hazard perception across driver collision groups (with error bars).

3.2.3. Differences between driving experience and collision groups in relation to their additional hazard responses

The number of extra hazard responses, beyond those responses identified within the predefined hazard window, was also calculated. The extra hazard response rate refers to the total number of additional responses recorded during the entire video that did not qualify as correct detections of the predefined hazard. The independent variables were driving experience and involvement in collisions and the dependent variable was the extra responses to the pre-defined hazards.

As the video clips varied in time, we modelled the responses as a Poisson count variable with an offset to account for the extra exposure for the duration of each clip (Baguley, 2012). The resulting multilevel generalized linear model, which included participant and clip as random factors, therefore estimated the extra hazard responses per minute (EHR/m) and was fitted using lme4 package. The deviance (likelihood ratio Chi Square χ^2) of the intercept only model was 7543.9 and decreased to 7540.3 for a model with main effects. The intercept only model ($\chi^2(3) = 7549.9$) appeared to be a worse fit than a model with all main effects (χ^2 (2) = 3.63, p = 0.16) and all two-way interactions (χ^2 (1) = 0.62, p = 0.43). No significant main effects and interactions were observed. No statistically significant differences were observed between experienced and novice drivers ($\chi^2(1) = 1.58$, p = 0.20), with similar mean scores for experienced drivers M = 2.07 and for novice drivers M = 2.19. Similarly, no significant differences were found between collision history groups ($\chi^2(1) = 3.22$, p = 0.07). Participants with no history of collisions had a mean score of M = 2.21, whereas those who had been involved in a collision had a mean score of M = 1.98 for additional hazard responses. The main effect of Collisions was not significant, indicating that participants with a history of collisions did not differ in click rate from those without such a history (IRR = 0.87, 95% CI [0.66, 1.15]). Similarly, Experience did not significantly predict click rate, with novices showing a comparable number of clicks to experienced drivers (IRR = 0.92, 95% CI [0.71, 1.19]).

3.3. *Hazard prediction test results*

The primary variable of interest in this analysis was the accuracy of drivers' ability to predict hazards. To examine differences in hazard prediction accuracy across various factors, a 2x2 factorial design was fitted using multilevel logistic regression with the R lme4 package (Bates et al., 2015). The independent variables consisted of the drivers' experience levels and their history of involvement in collisions, while the dependent variable was the accuracy of hazard prediction.

An intercept only model (with no predictors) estimated the SD of the participant random effect as 0.097 and the SD of the clip random effect as 0.729 indicating that only 9% variation at level 2 of the model is attributable to participants, with variability in clips accounting for the majority (72%) of level 2 variance. This indicates that a traditional ANOVA analysis, that treats variation between clips as zero, would substantially underestimate standard errors. The deviance (likelihood ratio Chi Square, $\Delta\chi^2$) for the intercept only model was 2228.9 and decreased substantially to 2192.2 for a model including main effects of experience and crashes. This improvement in model fit was statistically significant ($\Delta\chi^2$ (3) = 36.73, p < 0.001). However, the model with all two-way interactions was not statistically significant ($\Delta\chi^2$ (1) = 0.35, p = 0.55). The main effects model therefore appears to be the most informative. There were significant differences between the experienced and novice drivers in their ability to predict hazardous situations ($\Delta\chi^2$ (1) = 30.35, p > 0.001), (M = 78% vs. 48%) (see Fig. 9). The experienced drivers predicted significantly more hazards than the novices. No significant differences were found between drivers who had been involved in collisions and those who had not ($\Delta\chi^2$(1) = 0.02, p = 0.89), with mean accuracy rates of M = 65% and M = 64%, respectively (see Fig. 10). The effect of collision history was not significant, with odds of correct prediction for participants with a collision history being similar to those without (OR = 0.97, 95% CI [0.60, 1.56]). In contrast, novice drivers showed substantially lower odds of correct prediction compared with experienced drivers (OR = 0.26, 95% CI [0.17, 0.41]), indicating that novices were considerably less likely to make correct predictions. It is important to note that only 34 drivers reported being involved in collisions, compared to 72 experienced drivers (combining experienced and novice drivers in the hazard prediction task). Consequently, when experience was controlled for in the model, the analysis did not yield comparable differences, which limits our ability to draw conclusions regarding the impact of collisions. However, when examining the mean accuracy scores between **experienced drivers** who had been involved in collisions and those who had not, the scores were relatively similar at M = 75% and M = 73%, respectively.

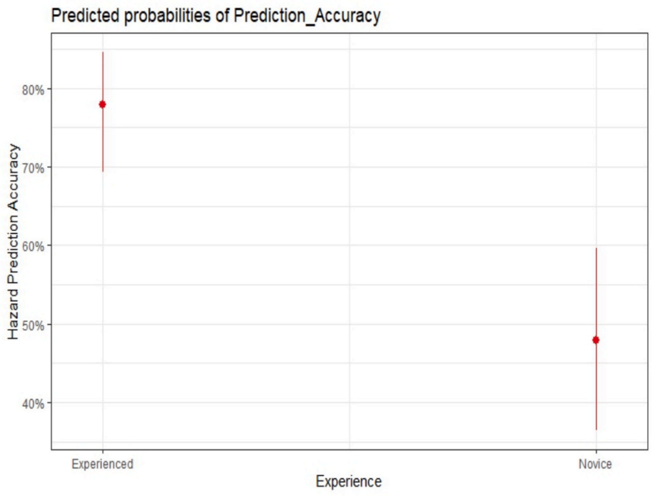


Fig. 9. Accuracy in hazard prediction across driving experience groups (with error bars).

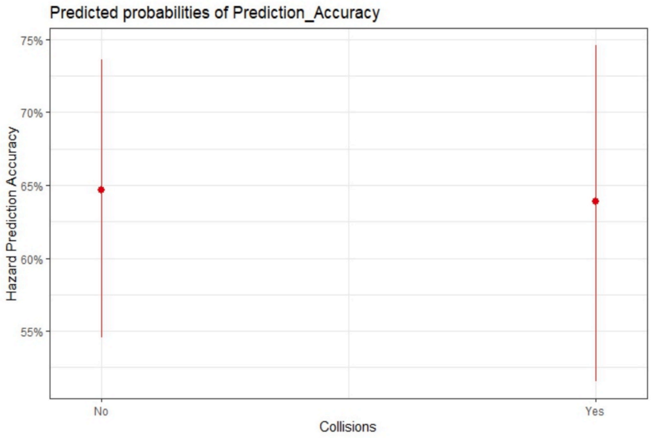


Fig. 10. Accuracy in hazard prediction across collision groups (with error bars).

3.4. *Comparisons of the two tasks*

Both tests were compared via logistic regression, with the factors being test type and driving experience, and the dependent variable – accuracy in hazard detection. The results indicated that there was no evidence that one test was more superior than the other in differentiating between the groups of drivers with varying levels of experience ($\Delta\chi^2$ (1) = 0.42, p = 0.52). The mean scores for hazard perception were *M* = 69% for experienced drivers and *M* = 49% for novice drivers, while for hazard prediction, the scores were *M* = 73% for experienced drivers and *M* = 48% for novices. Both tests demonstrated a robust ability to assess these skills. Additionally, no significant interaction was observed between experience and test type ($\Delta\chi^2$ (1) = 0.83, p = 0.36) (see Fig. 11). While experienced drivers outperformed novices in both tests (OR = 0.21, 95% CI [.13, 1.00]), no meaningful effects were observed for test type.

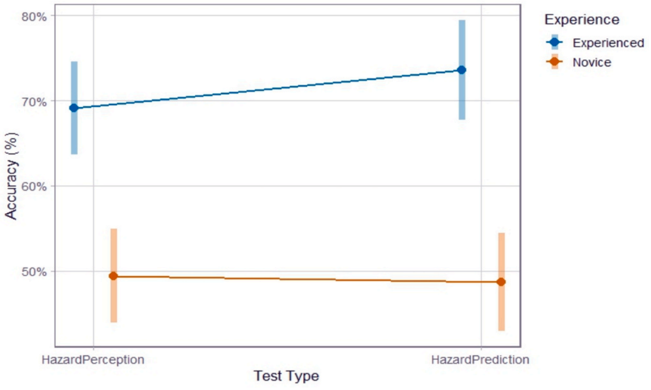


Fig. 11. Accuracy in perceiving hazards across test type and driving experience (with error bars). The central line within each box represents the median.

4. Discussion

The aim of the present study was to investigate whether a bespoke hazard perception (HP) test, specifically designed for the Czech driving context, could effectively distinguish between experienced and novice drivers. As evidence from prior studies has suggested that standardised HP tests may not transfer effectively across diverse driving contexts (Ventsislavova et al., 2019), two variants of the HP test were developed; a hazard perception test and a hazard prediction test. These tests reflect different levels of Situation Awareness (SA) theory (Endsley, 1995), assessing both drivers' ability to detect hazards in a timely manner and their ability to predict their development, and by developing both we were able to compare these two paradigms and evaluate which approach is better suited for the Czech driving context.

The findings revealed that experienced drivers significantly outperformed novices on both hazard perception and hazard prediction tests. Specifically, experienced drivers detected hazards more rapidly and demonstrated greater accuracy in both detecting and predicting hazardous situations. These findings are consistent with previous research, which has repeatedly shown that experienced drivers demonstrate superior situational awareness and hazard perception skills compared to novice drivers (e.g., (Hill et al., 2019; Wetton et al., 2011). Both hazard perception test variants developed for this study successfully distinguished between the two experience groups, producing promising results and supporting the potential integration of this form of assessment into the Czech driver licensing procedure. Notably, none of the participants had prior exposure to any hazard perception testing, yet experienced drivers consistently outperformed novices, indicating that the observed differences are attributable to driving experience rather than familiarity with the test format. Once again, the hazard perception test has demonstrated its capacity to differentiate between levels of driving expertise across both task variants. When considered alongside previous large-scale validation studies (e.g., (Ventsislavova et al., 2019; Ventsislavova et al., 2022; Ventsislavova et al., 2016) the present findings offer strong empirical support for the inclusion of hazard perception testing within national licensing systems, including those beyond the Czech Republic.

The observed differences in hazard perception and prediction accuracy between novice and experienced drivers are underpinned by cognitive and perceptual mechanisms related to attention allocation, situational awareness, and schema development. Experienced drivers typically possess more elaborate and refined mental models of traffic situations, enabling them to anticipate potential hazards even when visual cues are limited. This aligns with prior research indicating that driving experience enhances the ability to detect and interpret subtle environmental indicators related to the hazard (Crundall, 2016; Crundall et al., 2012). In contrast, novice drivers may exhibit narrower attentional focus and reduced capacity to integrate contextual information, resulting in poorer performance on hazard-related tasks. Similarly, individuals with a history of collision involvement may differ in their attentional control or threshold for threat appraisal, potentially reflecting either riskier driving behaviour or deficits in processing dynamic traffic information (e.g., (Wetton et al., 2011; Castro et al., 2021).

In this line, the findings also revealed statistically significant differences between collision-involved and collision-free drivers, with the former group exhibiting lower accuracy and slower response times in detecting hazards. While these findings should be interpreted with caution due to the unequal sample sizes between collision groups, the hazard perception test appeared more effective than the hazard prediction test in distinguishing between drivers based on their collision history. While further research is needed to better understand the relationship between hazard-related performance and crash involvement, the current results seem promising. These findings corroborate existing evidence indicating that superior hazard perception performance is associated with reduced crash risk across a range of driver populations (e.g., (Boufous et al., 2011; Drummond, 2000).

Nonetheless, the hazard prediction test did not distinguish between collision-involved and collision-free drivers. To the authors' knowledge, only one study has attempted to link self-reported collisions with hazard perception performance, conducted by Horswill et al. (Horswill et al., 2020). Their results indicated that drivers who reported no crashes in the previous two years made significantly more valid predictions in the hazard prediction test than those who reported crashes, however, this finding is constrained by the use of only two video clips to measure hazard perception performance. Given the mixed evidence, additional research is warranted to determine whether hazard prediction skill is a valid predictor of crash risk.

It should also be noted that experienced drivers achieved relatively high scores on the hazard prediction test, a finding that is not without precedent. While some previous studies have shown the hazard prediction test to be harder than the hazard perception test for both novice and experienced drivers (e.g., (Ventsislavova et al., 2019; 2022), other studies have reported comparable levels of performance among experienced drivers. For example, Ventsislavova and Crundall (Ventsislavova and Crundall, 2018) reported an accuracy of 72% in a study examining three different occlusion points in a hazard prediction task. Their results indicated that shorter temporal gaps between the hazard precursor and the subsequent development of the hazard were associated with higher prediction accuracy. This effect is likely due to the unfolding nature of hazard precursors, which may also have contributed to the present findings. Similarly, Crundall and Kroll (Crundall and Kroll, 2018) found that low-risk, experienced fire-appliance drivers achieved an average accuracy of 73% on a hazard prediction test. Eye-tracking data revealed that these drivers fixated hazard precursors significantly more often than higher-risk drivers, indicating greater attentional prioritisation of locations from which hazards were likely to emerge. These findings support the theoretical rationale underpinning hazard prediction testing, namely that safer drivers are more effective at identifying, prioritising, and monitoring potential precursors, and are therefore more likely to be attending to the correct location at the critical moment. Importantly, the hazard prediction test in the current study still successfully differentiated between experienced and novice drivers demonstrating that experienced drivers were more likely to prioritise areas where hazards were anticipated to emerge.

The findings of the present study also indicated that experienced and novice drivers produced a comparable number of additional hazard responses, suggesting similar levels of general vigilance. These results align with previous cross-cultural research comparing hazard perception performance in China, Spain and the UK, which reported no significant differences between experienced and novice drivers but found differences across nationalities (Ventsislavova et al., 2019). However, the significant experience group differences observed in hazard perception performance suggest that experienced drivers were more accurate at detecting and responding to the pre-defined hazards. This implies that their performance advantage did not stem from heightened general responsiveness, but rather from more efficient and accurate hazard discrimination. In contrast, novice drivers may have produced additional responses due to uncertainty yet failed to demonstrate the level of cognitive filtering and prioritisation of relevant cues observed in experienced drivers. While SA does not inherently involve decision-making, it plays a pivotal role in guiding the selection of appropriate behavioural responses that enhance driving safety.

It was somewhat surprising that the hazard prediction test did not distinguish between collision-involved and collision-free drivers. One possible explanation for the current pattern of results relates to the nature of the response required. For example, Ventsislavova et al. (Ventsislavova et al., 2016) reported no significant differences in hazard prediction between offender and non-offender drivers, however, differences emerged in post-perceptual decision-making, with a significantly greater proportion of non-offender drivers indicating that they would perform an evasive manoeuvre to avoid hazards than offender drivers. A further consideration is that any measure of collision involvement may include collisions that could not have been mitigated by hazard perception skill, therefore weakening the potential link between crashes and hazard perception score, especially in smaller samples (Af Wåhlberg, 2018). This may explain why the collision effect is smaller than the experiential effect in the hazard perception analysis and absent in the hazard prediction analysis.

Notably, research conducted in countries outside the United Kingdom and Australia has often shown hazard prediction tasks to be more effective than perception tasks in distinguishing between driver groups (e.g., (Ventsislavova et al., 2019; Lim et al., 2013). These findings have been attributed to differences in criterion bias – the internal threshold individuals use to judge whether a situation qualifies as hazardous. In high-risk driving environments, such as those in Malaysia and China, drivers may develop a higher threshold for defining an event as hazardous. This could potentially influence their responses on perception-based tasks, as while such drivers may detect and even fixate on a hazard, they may delay overt responses until the situation crosses their personal threshold for action (e.g. (Lim et al., 2013). However, the current study demonstrated that both test types were effective in distinguishing between levels of driving experience, following the same test development protocol outlined in Ventsislavova et al. (Ventsislavova et al., 2019; Ventsislavova et al., 2022). These findings suggest that the Czech driving context may be more closely aligned with those of countries such as the United Kingdom and Australia, where both hazard perception and hazard prediction tasks have proven effective. Such alignment may reflect underlying similarities in road infrastructure, traffic legislation, and common driving practices.

Hazard perception and prediction tasks are not merely visual recognition tasks. Rather, they draw upon higher-order cognitive processes such as working memory, predictive reasoning, and decision-making under conditions of uncertainty (Horswill et al., 2004; Pradhan et al., 2017). However, even a good level of situational awareness does not guarantee collision-free driving, as some drivers with relatively poor situational awareness may still avoid collisions due to other compensatory factors. Conversely, drivers with good situational awareness may nonetheless be involved in crashes if they fail to act on the information effectively or if other risk factors are present. Collision involvement is typically the result of a complex interplay of variables, including risk proneness, demographic factors, and environmental conditions, all of which must be considered in interpreting hazard-related performance (e.g., (Castro et al., 2021; Ventsislavova et al., 2021). The willingness to engage in risky driving behaviour reflects a driver's propensity to adjust their actions according to the level of risk they are personally willing to accept (Hergovich et al., 2007). Importantly, even when drivers accurately perceive a hazardous situation, they may still underestimate the degree of risk it presents, potentially leading to unsafe decisions (Ventsislavova et al., 2016). Thus, differences between collision-involved and collision-free drivers may be more pronounced in post-perceptual processes such as risk calibration and decision-making. As situational awareness serves as a precursor to decision-making in driving contexts (Endsley, 2021), the hazard prediction task does not encompass the full range of post-perceptual processes involved in hazard avoidance, such as hazard appraisal and behavioural response selection, nor is it intended to do so. Rather, its primary function is to assess drivers' awareness of the driving environment and their ability to anticipate how a situation is likely to unfold, without conflating hazard prediction with hazard appraisal mechanisms (e.g. (Pradhan et al., 2017).

More specifically, the hazard prediction task primarily assesses drivers' ability to project how a situation is likely to unfold under conditions of reduced time pressure. In contrast, the hazard perception task requires timely situational appraisal and sensitivity to temporal dynamics as hazards emerge in real time. Slower response latencies among collision-involved drivers therefore may suggest inefficiencies not in hazard knowledge per se, but in determining when a developing situation has reached a level of criticality requiring immediate action. Although both hazard prediction and hazard perception rely on early perceptual–attentional processing, the divergence between groups may arise at a later stage, specifically at the point at which drivers must translate awareness into timely action. Collision-involved drivers may thus demonstrate intact predictive capability while exhibiting deficiencies in the timing or calibration of their behavioural responses.

Under real-world time pressure, even modest delays in judgement or response initiation compress the window available for downstream processes such as risk appraisal and motor execution, thus reducing safety margins and increasing collision likelihood. For that reason, drivers with higher risk tolerance may adopt a later decision threshold, requiring greater hazard imminence before initiating evasive action. In this account, the perceptual representation of the hazard is preserved, but the criterion for responding is shifted, resulting in delayed intervention despite intact predictive awareness. Crashes, therefore, may occur not because hazards are unrecognised, but because action is initiated too late relative to environmental demands. An alternative, though not mutually exclusive, explanation concerns perceptual discrimination under time pressure. Collision-involved drivers may experience difficulty distinguishing diagnostic information (signal) from competing background stimuli (noise), leading to slower evidence accumulation and delayed hazard detection. Such inefficiencies would be more likely to manifest in real-time perception tasks than in prediction tasks administered under reduced temporal constraints.

Both elevated response thresholds and signal–noise discrimination difficulties have been discussed in the literature [e.g., 69, 61, 70]. However, further research employing methodologies capable of isolating decision criterion placement and perceptual sensitivity would be required to determine whether collision involvement is more strongly associated with altered response thresholds, reduced perceptual discrimination efficiency, or an interaction between the two, including potential cross-cultural variation.

When the two tests were compared, the results indicated that neither demonstrated superiority in distinguishing between experienced and novice drivers. Consequently, for optimal driver assessment, training, and hazard awareness (particularly among novice drivers), we recommend that both hazard perception and hazard prediction tests be incorporated into the Czech licensing system. Each test offers distinct yet complementary insights into drivers' hazard-related competencies. Assessing these two key sub-processes separately allows for a more comprehensive understanding of the complex cognitive and behavioural demands of safe driving. The key value of hazard perception tasks lies in its capacity to assess the speed with which drivers detect and respond to emerging hazards, whereas hazard prediction tasks are instrumental in evaluating drivers' ability to understand and anticipate potentially hazardous scenarios, even when such scenarios are not perceived as hazardous by drivers. This distinction is particularly pertinent in high-risk driving environments, and notably in the context of the present study, where collision-involved drivers appeared to comprehend the hazard yet failed to respond to it appropriately. Hazard perception extends beyond the mere detection of a hazard. It involves a broader cognitive process in which sensory information is interpreted and integrated with prior knowledge and contextual cues to construct a mental model of the driving environment (Endsley, 1995; Bellet et al.; Endsley, 1995). This process ultimately supports the recognition and anticipation of potential hazards. The inclusion of a holistic hazard perception component within national licensing systems is therefore critical for enhancing situational awareness and encouraging timely and effective responses to hazards.

While novice drivers may exhibit rapid reactions to hazards, this does not necessarily indicate well-developed SA. Their responses may reflect reflexive behaviour rather than informed anticipation of hazardous scenarios, and such reactions may not consistently result in safe outcomes. Therefore, hazard perception should be approached holistically, incorporating not only the detection of hazards but also the assessment of drivers' anticipatory skills. In this regard, situational awareness levels 2 (comprehension; e.g., “This object may interfere with my trajectory”) and 3 (projection; e.g., “This object will cut my path”) are generally more informative indicators of driver competence than level 1 (perception), which serves primarily as a necessary foundation for higher-level processing (de Zwart et al., 2025). While the ideal goal is to achieve complete SA, in practice this is rarely attainable, instead, SA is conceptualised as a probabilistic construct, varying across individuals and situations (Endsley, 2021; Endsley et al., 2000). Not all elements in the driving environment carry equal relevance for hazard monitoring, and effective hazard perception relies on the driver’s ability to prioritise critical cues. A failure to accurately identify elements that pose an immediate threat indicates an underdeveloped level of SA. Crucially, the act of observing an element does not, in itself, imply comprehension of its significance; meaningful interpretation depends on context-specific expectations shaped by prior experience and knowledge (e.g., (Bellet et al.). A deeper understanding of how the levels of situational awareness operate within dynamic driving contexts, and how they relate to risk assessment and decision-making, is essential for clarifying the relationship between pre-perceptual and post-perceptual processes in hazard perception.

4.1. Limitations

It should be noted that the collision groups in this study were unequal, with the number of collision-free drivers approximately three times greater than those involved in collisions. This imbalance may have influenced the statistical power of the analyses, and as such, the results should be interpreted with caution. While multilevel modelling is generally robust to unbalanced data and designed to accommodate variability at different levels, substantial group size disparities can still influence the precision of parameter estimates and standard errors. Specifically, the smaller group (collision-involved) may have contributed less information to the model, potentially reducing the power to detect effects in this group for the hazard prediction test. This imbalance may partly account for the mixed pattern of significant and non-significant findings, even when the direction of effects appeared consistent. Future studies should prioritise focusing specifically on collision groups to strengthen the robustness and interpretability of group effects and strive to measure only those collisions that could be plausibly due to poor hazard skills of our sample.

As is common in experimental driving research, reliance on voluntary participation may limit the generalisability of the findings. Future research would therefore benefit from replication using stratified or population-based sampling approaches. In addition, due to the very low accumulated mileage and the largely supervised nature of learners’ driving experience, self-reported collisions among learner drivers may differ qualitatively from those experienced by fully licensed drivers in unsupervised traffic conditions. As detailed information regarding collision circumstances (e.g., supervision, driver role, or severity) was not systematically collected, comparisons involving collision history should be interpreted with caution. Future research should aim to record and examine such contextual information.

Finally, one further possible limitation lies in the design of the hazard prediction test, as participants have time between occlusion and response to fully consider their answer, arguably overestimating their ability to make decisions ‘in the moment’. However, we do not argue that hazard prediction reflects in-the-moment decisions. It does, however, reflect what visual information we have obtained in the immediate moment prior to occlusion. While it is possible that highly experienced drivers might be able to infer the answer based on information obtained that was not directly pertinent to the imminent hazard, previous research has shown the test to be sensitive to what participants fixate immediately prior to fixation (Crundall and Kroll, 2018). This suggests that the visual information obtained prior to occlusion is vital to success in this task. While we acknowledge the possibility that post-perceptual decision making will influence accuracy, this does not detract from the many studies that have demonstrated the success of hazard prediction tests to differentiate between driver groups.

4.2. Implications

The present study provides evidence supporting the implementation of the newly developed hazard perception and prediction tests within the Czech licensing system. Importantly, this is not the first study to successfully adapt hazard perception testing outside the UK and Australia, where the approach was originally developed. Research conducted in China, Spain, the UK (Ventsislavova et al., 2019; Castro et al., 2014; Crundall et al., 2021), Lithuania (Endriulaitienė et al., 2022), Israel (Ventsislavova et al., 2022), Malaysia (Lim et al., 2013; 2014), Germany (Malone and Brünken, 2016), and other countries has demonstrated that standardised guidelines can be used to develop valid and contextually appropriate hazard perception assessments. Such efforts support consistent training and evaluation across diverse driving environments. The HP test has already been incorporated into the official driving licence assessments in several countries, including some Australian states, the United Kingdom, and the Netherlands. While a number of other countries, such as Spain, Germany, or Canada, have expressed interest in adopting HP testing, it has yet to be formally implemented in many countries. The introduction of HP testing presents considerable challenges, as several key factors must be addressed. These include the legal and regulatory frameworks, existing infrastructure, alignment with national crash data, and integration within current training and educational systems. Moreover, any HP test must be compatible with the specific licensing system of each country. This complexity is further compounded by the fact that each country presents a distinct set of hazardous driving scenarios, shaped by its own legal frameworks and social norms. Therefore, in order to implement a standardised HP assessment globally, it is necessary to develop culturally and contextually specific sets of hazard clips that reflect both common and critical hazardous situations within each country. Simultaneously, the methodology including test design, instructions, and format must be adapted and validated to ensure that it reliably discriminates between novice and experienced drivers. Previous studies have provided detailed guidelines to facilitate the development of such a standardised methodology (e.g., (Ventsislavova et al., 2019; Ventsislavova and Crundall, 2018; Ventsislavova et al., 2022). In line with previous research, the current study employed established protocols to maintain methodological consistency and to validate the measures for implementation in the Czech licensing system, thereby confirming the tests’ robustness and suitability for this purpose.

5. Conclusion

This research adds to the substantial body of evidence that hazard perception testing can effectively differentiate between drivers of varying experience levels, supporting its potential inclusion in the official licensing system of the Czech Republic. In the present study, both tests once again demonstrated their capacity to distinguish between experienced and novice drivers, showing robustness and sensitivity in detecting differences in situational awareness across experiential groups. However, only the hazard perception test was able to distinguish between drivers with and without prior collision involvement. These findings highlight the critical role of hazard perception in driving, as it involves the cognitive integration of sensory information, prior knowledge, and contextual cues to build an accurate mental representation of the driving environment. Accordingly, the integration of a robust hazard perception component into the Czech national licensing system is essential for enhancing situational awareness and enabling timely, effective responses to potential hazards, ultimately supporting a reduction in collision risk.

CRediT authorship contribution statement

Petya Ventsislavova: Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Matus Sucha:** Writing – review & editing, Writing – original draft, Resources, Investigation, Funding acquisition, Conceptualization. **Lydia Harrison:** Writing – review & editing, Writing – original draft, Visualization, Resources, Project administration, Formal analysis, Data curation. **Jiri Novotny:** Investigation, Funding acquisition. **David Crundall:** Writing – review & editing, Methodology, Conceptualization.

Declaration of competing interest

Acknowledgement

We would like to thank Mikuláš Toman for assisting with the filming of the driving footage, which was subsequently edited into the hazard clips.

Funding

This work was supported by the Technology Agency of the Czech Republic [CK03000009].

Data availability

Data will be made available on request.

References

European Transport Safety Council, (2025). Ranking EU progress on road safety: 19th Road safety performance index (PIN) report. https://etsc.eu/wp-content/uploads/ETSC-2025-Annual-PIN-Report-DIGITAL-V2.pdf.

World Health Organization, 2018. Global status report on road safety 2018. World Health Organization.

World Health Organization, 2013. Global status report on road safety 2013: supporting a decade of action. World Health Organization.

United Nations, 2010. Resolutions adopted by the general assembly on 2 March 2010. United Nations. https://www.undocs.org/A/RES/64/255.

CINEA, 2022. EU road safety: Towards "vision zero". European Union. https://cinea.ec.europa.eu/system/files/2023-02/H2020%20Transport-Road%20Safety%202022-web.pdf.

International Transport Forum, 2022. Road safety annual report 2022. OECD Publishing, Paris.

White, M.J., Cunningham, L.C., Titchener, K., 2011. Young drivers' optimism bias for accident risk and driving skill: accountability and insight experience manipulations. Accid. Anal. Prevent. 43, 1309–1315. https://doi.org/10.1016/j.aap.2011.01.013.

European Transport Safety Counci, 2019. Road safety priorities for the EU 2020 – 2030. https://etsc.eu/wp-content/uploads/2019-07-New-EP-briefing-1.pdf.

Wells, P., Tong, S., Sexton, B., Grayson, G., Jones, E., 2008. Cohort II: a study of learner and new drivers: Volume 1 (Road Safety Research Report No. 81). London: Department for Transport.

Ventsislavova, P., Crundall, D., Baguley, T., Castro, C., Gugliotta, A., Garcia-Fernandez, P., Zhang, W., Ba, Y., Li, Q., 2019. A comparison of hazard perception and hazard prediction tests across China, Spain and the UK. Accid. Anal. Prevent. 122, 268–286. https://doi.org/10.1016/j.aap.2018.10.010.

Endsley, M.R., 1995. Toward a theory of Situation Awareness in dynamic systems. Human Factors: J. Hum. Factors Ergonom. Soc. 37 (1), 32–64. https://doi.org/10.1518/001872095779049543.

Crundall, D., 2016. Hazard prediction discriminates between novice and experienced drivers. Accid. Anal. Prevent. 86, 47–58. https://doi.org/10.1016/j.aap.2015.10.006.

Horswill, M.S., Taylor, K., Newnam, S., Wetton, M., Hill, A., 2013. Even highly experienced drivers benefit from a brief hazard perception training intervention. Accid. Anal. Prevent. 52, 100–110. https://doi.org/10.1016/j.aap.2012.12.014.

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

Horswill, M.S., Hill, A., Wetton, M., 2015. Can a video-based hazard perception test used for driver licensing predict crash involvement? Accid. Anal. Prevent. 82 (23), 213–219. https://doi.org/10.1016/j.aap.2015.05.019.

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. Accid. Anal. Prevent. 122, 207–214. https://doi.org/10.1016/j.aap.2018.08.030.

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. Traff. Injury Prevent. 12 (5), 475–482. https://doi.org/10.1080/15389588.2011.591856.

Crundall, D., Chapman, P., Trawley, S., Collins, L., van Loon, E., Andrews, B., Underwood, G., 2012. Some hazards are more attractive than others: drivers of varying experience respond differently to different types of hazard. Accid. Anal. Prevent. 45, 600–609. https://doi.org/10.1016/j.aap.2011.09.049.

Horswill, M.S., McKenna, F.P., 1999. The development, validation, and application of a video-based technique for measuring an everyday risk-taking behavior: drivers' speed choice. J. Appl. Psychol. 84 (6), 977–985. https://doi.org/10.1037/0021-9010.84.6.977.

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

Drummond, A.E., 2000. Paradigm lost! paradise gained? An Australian perspective on novice driver safety. Proceedings of the Novice Drivers Conference.

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. Accid. Anal. Prevent. 43 (4), 1464–1471. https://doi.org/10.1016/j.aap.2011.02.024.

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

Horswill, M.S., Anstey, K.J., Hatherly, C.G., Wood, J.M., 2010. The crash involvement of older drivers is associated with their hazard perception latencies. J. Int. Neuropsychol. Soc. 16 (5), 939–944. https://doi.org/10.1017/s135561771000055x.

Rosenbloom, T., Perlman, A., Pereg, A., 2011. Hazard perception of motorcyclists and car drivers. Accid. Anal. Prevent. 43 (3), 601–604. https://doi.org/10.1016/j.aap.2010.08.005.

Ross, R.W.J., Cordazzo, S.T.D., Scialfa, C.T., 2014. Predicting on-road driving performance and safety in healthy older adults. J. Saf. Res. 51, 73–80. https://doi.org/10.1016/j.jsr.2014.09.005.

Wood, J.M., Horswill, M.S., Lacherez, P.F., Anstey, K.J., 2013. Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. Accid. Anal. Prevent. 50, 1161–1168. https://doi.org/10.1016/j.aap.2012.09.009.

Mills, K.L., Hall, R.D., McDonald, M., Rolls, G.W.P., 1998, The effects of hazard perception training on the development of novice driver skills, In: Roadsafe 96 International Conference, Influences Affecting Road User Behaviour, London, United Kingdom.

Wetton, M.A., Horswill, M.S., Hatherly, C., Wood, J.M., Pachana, N.A., Anstey, K.J., 2010. The development and validation of two complementary measures of drivers' hazard perception ability. Accid. Anal. Prevent. 42, 1232–1239. https://doi.org/10.1016/j.aap.2010.01.017.

Sagberg, F., Bjørnskau, T., 2006. Hazard perception and driving experience among novice drivers. Accid. Anal. Prevent. 38, 407–414. https://doi.org/10.1016/j.aap.2005.10.014.

Yeung, J.S., Wong, Y.D., 2015. Effects of driver age and experience in abrupt-onset hazards. Accid. Anal. Prevent. 78, 110–117. https://doi.org/10.1016/j.aap.2015.02.024.

Chapman, P., Underwood, G., 1998. Visual search of driving situations: danger and experience. Perception 27, 951–964. https://doi.org/10.1068/p270951.

Underwood, G., Phelps, N., Wright, C., van Loon, E., Galpin, A., 2005. Eye fixation scanpaths of younger and older drivers in a hazard perception task. Ophthal. Physiol. Opt. 25 (4), 285–379. https://doi.org/10.1111/j.1475-1313.2005.00290.x.

Horswill, M.S., Hill, A., 2021. Driver's hazard perception skill. Int. Encyclop. Transp. 2021, 151–157. https://doi.org/10.1016/B978-0-08-102671-7.10678-5.

McCartt, A.T., Mayhew, D.R., Braitman, K.A., Ferguson, S.A., Simpson, H.M., 2009. Effects of age and experience on young driver crashes: Review of recent literature. Traff. Injury Prevent. 10 (3), 209–219. https://doi.org/10.1080/15389580802677807.

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

Endriulaitienė, A., Šeibokaitė, L., Markšaitytė, R., Slavinskienė, J., Crundall, D., Ventsislavova, P., 2022. Correlations among self-report, static image, and video-based hazard perception assessments: the validity of a new Lithuanian hazard prediction test. Accid. Anal. Prevent. 173, 106716. https://doi.org/10.1016/j.aap.2022.106716.

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

Konstantopoulos, P., Chapman, P., Crundall, D., 2010. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accid. Anal. Prevent. 42, 827–834. https://doi.org/10.1016/j.aap.2009.09.022.

Borowsky, A., Shinar, D., Oron-Gilad, T., 2010. Age, skill and hazard perception in driving. Accid. Anal. Prevent. 42 (4), 1240–1249. https://doi.org/10.1016/j.aap.2010.02.001.

Ventsislavova, P., Crundall, D., 2018. The hazard prediction test: a comparison of free response and multiple-choice formats. Saf. Sci. 109, 246–255. https://doi.org/10.1016/j.ssci.2018.06.004.

Sivak, M., Soler, J., Tränkle, U., Spangnhol, J.M., 1988. Cross-cultural differences in driver risk-perception. Accid. Anal. Prevent. 21 (4), 355–362. https://doi.org/10.1016/0001-4575(89)90026-2.

Bazilinskyy, P., Eisma, Y.B., Dodou, D., de Winter, J.C.F., 2020. Risk perception: a study using dashcam videos and participants from different world regions. Traff. Injury Prevent. 21 (6), 347–353. https://doi.org/10.1080/15389588.2020.1762871.

Caparelli-Daquer, E., Santana, T., Cordazzo, S., Cordazzo, H., Scialfa, C.T., 2017. Hazard perception test (HPT): A pilot study in Brazil. In: Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (72 – 78). Public policy center, University of Iowa. .

Lim, P., Sheppard, E., Crundall, D., 2013. Cross-cultural effects on drivers' hazard perception. Transp. Res. Part F: Traff. Psychol. Behav. 21, 194–206. https://doi.org/10.1016/j.trf.2013.09.016.

Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: Bridging selective attention and working memory. Trends Cognit. Sci. 16 (2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014.

Pradhan, A.K., Crundall, D., 2017. Hazard avoidance in young novice drivers: definitions and a framework. In: Fisher, D.L., Caird, J., Horrey, W., Trick, L. (Eds.), Handbook of Teen and Novice Drivers, (1st Ed.). CRC Press.

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.

Lim, P., Sheppard, E., Crundall, D., 2014. A predictive hazard perception paradigm differentiates driving experience cross-culturally. Transp. Res. Part F: Traff. Psychol. Behav. 26, 210–217. https://doi.org/10.1016/j.trf.2014.07.010.

Ventsislavova, P., Rosenbloom, T., Leunissen, J., Spivak, Y., Crundall, D., 2022. An online hazard prediction test demonstrates differences in the ability to identify hazardous situations between different driving groups. Ergonomics 65 (8), 1119–1137. https://doi.org/10.1080/00140139.2021.2016999.

Horswill, M.S., Hill, A., Jackson, T., 2020. Scores on a new hazard prediction test are associated with both driver experience and crash involvement. Transp. Res. Part F: Traff. Psychol. Behav. 71, 98–109. https://doi.org/10.1016/j.trf.2020.03.016.

Manley, H., Paisarnsrisomsuk, N., Hill, A., Horswill, M.S., 2020. The development and validation of a hazard perception test for Thai drivers. Transp. Res. Part F 71, 229–237. https://doi.org/10.1016/j.trf.2020.04.011.

Sun, L., Li, S., Chang, R., 2019. Development and validity of Chinese hazard perception test. In: Long, S., Dhillon, B.S. (Eds.), Man-Machine-Environment System Engineering. Springer, pp. 257–262. https://doi.org/10.1007/978-981-13-2481-9_30.

European Commission, Directorate-General for Mobility and Transport, 2022. EU transport in figures: statistical pocketbook 2022, Publications Office of the European Union. .

European Commission, 2021. Facts and figures young people. European road safety observatory. Brussels, European Commission, Directorate General for Transport. https://road-safety.transport.ec.europa.eu/system/files/2022-01/F%26F_young_people_20211221.pdf.

BESIP, 2021. Czech road traffic safety strategy 2021 – 2030. Prague, Road safety department, Ministry of transport, https://besip.cz/getattachment/Pro-odborniky/Narodni-strategie-BESIP/Aktualni-strategie/Czech-Road-Traffic-Safety-Strategy-2021-30_11-11.pdf?lang=cs-CZ.

Baguley, T., 2012. Serious stats: a guide to advanced statistics for the behavioural sciences. Palgrave Macmillan, London.

R Core Team, 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Christensen, R.H.B., 2018. Ordinal - regression models for ordinal data. R Package Version 2018. pp. 4–19.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1–48.

Ventsislavova, P., Gugliotta, A., Peña-Suarez, E., Garcia-Fernandez, P., Eisman, E., Crundall, D., Castro, C., 2016. What happens when drivers face hazards on the road? Accid. Anal. Prevent. 91, 43–54. https://doi.org/10.1016/j.aap.2016.02.013.

Castro, C., Ventsislavova, P., García-Fernández, P., Crundall, D., 2021. Risky decision-making and hazard prediction are negatively related and could be assessed independently using driving footage. Psychol. Res. Behav. Manag. 14, 857–876. https://doi.org/10.2147/prbm.s305979.

Crundall, D., Kroll, V., 2018. Prediction and perception of hazards in professional drivers: does hazard perception skill differ between safe and less-safe fire-appliance drivers? Accid. Anal. Prevent. 121 (6), 335–346. https://doi.org/10.1016/j.aap.2018.05.013.

Af Wåhlberg, A.E., 2018. The effect of driver improvement interventions on crash involvement; has it been under-estimated? Transp. Res. Part F: Traff. Psychol. Behav. 54, 349–356. https://doi.org/10.1016/j.trf.2018.02.027.

Ventsislavova, P., Crundall, D., Garcia-Fernandez, P., Castro, C., 2021. Assessing willingness to engage in risky driving behaviour using naturalistic driving footage: the role of age and gender. Int. J. Environ. Res. Public Health 18 (19), 10227. https://doi.org/10.3390/ijerph181910227.

Hergovich, A., Arendasy, M.E., Sommer, M., Bognar, B., 2007. The vienna risk-taking test - traffic. J. Individ. Differ. 28 (4), 198–204. https://doi.org/10.1027/1614-0001.28.4.198.

Endsley, M.R., 2021. A systematic review and meta-analysis of direct objective measures of situation awareness: a comparison of SAGAT and SPAM. Human Factors: J. Hum. Factors Ergon. Soc. 63 (1), 124–150. https://doi.org/10.1177/0018720819875376.

Bellet, T., Bailly-Asuni, B., Mayenobe, P., Banet, A., A theoretical and methodological framework for studying and modelling drivers' mental representations. Safety Science, 47(9), 1205 – 1221. https://doi.org/10.1016/j.ssci.2009.03.014.

Endsley, M.R., 1995. Measurement of situation awareness in dynamic systems. Human Factors: J. Hum. Factors Ergon. Soc. 37 (1), 65–84. https://doi.org/10.1518/001872095779049499.

de Zwart, R., Jansen, R.J., Bolstad, C., Endsley, M.R., Ventsislavova, P., de Winter, J., Young, M.S., 2025. When is more actually better? Expert opinions on assessment of situation awareness in relation to safe driving. Transp. Res. Part F: Traff. Psychol. Behav. 108, 54–72. https://doi.org/10.1016/j.trf.2024.11.011.

Endsley, M.R., 2000. Theoretical underpinnings of situation awareness: a critical review. In: Endsley, M.R., Garland, D.J. (Eds.), Situation Awareness Analysis and Measurement, (1st ed.). CRC Press. https://doi.org/10.1201/b12461.

Castro, C., Padilla, J.L., Roca, J., Benítez, I., García-Fernández, P., Estévez, B., López-Ramón, M.F., Crundall, D., 2014. Development and validation of the spanish hazard perception test. Traff. Injury Prevent. 15 (8), 817–826. https://doi.org/10.1080/15389588.2013.879125.

Crundall, D., van Loon, E., Baguley, T., Kroll, V., 2021. A novel driving assessment combining hazard perception, hazard prediction and theory questions. Accid. Anal. Prevent. 149, 105847. https://doi.org/10.1016/j.aap.2020.105847.

Malone, S., Brünken, R., 2016. The role of ecological validity in hazard perception assessment. Transp. Res. Part F: Traff. Psychol. Behav. 40, 91–103. https://doi.org/10.1016/j.trf.2016.04.008.