Cognitive predictors of hazard perception in young provisional drivers

Stephanie C. Goodhew[1], Mark Edwards[1], Andrew Hill[2,3], Mark S. Horswill[2]

[1]*School of Medicine and Psychology, The Australian National University, Australia*

[2]*School of Psychology, The University of Queensland, Australia*

[3]*Minerals Industry Safety and Health Centre, Sustainable Minerals Institute, The University of Queensland, Australia*

Corresponding author: Stephanie C. Goodhew

Address: Room 122, Psychology Building (building 39)

The Australian National University, Canberra, 2601

Email: Stephanie.Goodhew@anu.edu.au

Running head: Cognitive predictors of young drivers' crash risk

Abstract

Young drivers are at elevated crash risk compared with older age groups. Identifying the cognitive factors that can predict a driver's crash risk within this cohort is important because it can identify novel processes to target for cognitive training programs to improve road safety, and it can also help to identify those drivers who could benefit most from such interventions. We report a study in which 147 drivers aged between 17 and 27 years with a provisional Australian driver's license completed a video-based test of hazard perception, which has shown to be associated with indices of real-world driving behaviour and crash risk, and a suite of cognitive predictors selected from executive function theory and previous applied work in other driver populations. Results showed that scores on the Cognitive Failures Questionnaire 2.0, which gauges the frequency of errors in attention and memory in everyday life (e.g., forgetting appointments), were associated with hazard perception response times, such that drivers who experienced more frequent cognitive failures were slower to detect hazards. In contrast, none of the performance-based cognitive measures were associated with hazard perception response time. This suggests that how young drivers use their cognition in everyday life may be especially important to understanding crash risk.

Keywords: driving; young drivers; hazard perception; Useful Field of View; cognitive failures; executive function; cognition

Cognitive predictors of hazard perception and crash risk in young drivers

As a cohort, young drivers are at elevated crash risk (Thompson et al., 2018) but, within this cohort, there are individual differences in crash risk. Cognitive factors are crucial to driving, and identifying the cognitive processes associated with crash risk for young drivers is important for two key reasons. First, it can help to identify the relevant processes to target for interventions such as training. For example, if it is established that a given cognitive process is associated with safer driving (reduced crash risk), then it can inform interventions designed to scaffold and enhance this cognitive process. Second, it can help to identify the individuals within the young driver cohort who are most in need of such interventions to reduce their crash risk.

Previous research shows that both driving style factors (e.g., speed choice) and driving skill (e.g., hazard perception) contribute to crash risk (e.g., Elander et al., 1993; Hill et al., 2023; Horswill, 2016). Previous research has documented associations between cognitive processes and driving *style* factors in young drivers, such as speed choice in a simulated drive (Di Meco et al., 2021; Ledger et al., 2019; Zicat et al., 2018). However, less research has focussed on cognitive predictors of driving *skill* in young drivers. *Hazard perception* is a key aspect of driving skill that is predictive of crash risk (Horswill, 2016). Hazard perception can be thought of as drivers' situation awareness in traffic – their capacity to predict dangerous instances that may require their action (e.g., braking) to avoid a collision (Hill et al., 2019; Horswill, 2016; Horswill & McKenna, 2004). Given the cognitively demanding nature of hazard perception, there is good reason to predict that cognitive factors may be implicated in hazard perception. The goal of the present research was therefore to identify potential cognitive predictors of hazard perception skill in young drivers. The specific cognitive predictors considered were identified via: (a) the application of executive function theory to the demands of hazard perception and driving, and (b) from cognitive predictors that have been found to predict crash risk in other populations (e.g., older drivers). The rationale for each predictor is explained in detail below.

Cognitive predictors

Executive Functions. Executive functions are the domain-general cognitive processes that enable goal-directed action, especially where this requires cognitive flexibility, creating and revising mental representations, and resisting salient but task-irrelevant stimuli or habitual responses (Miyake et al., 2000). Hazard perception can be conceptualised as requiring executive functions, since it entails creating and frequently updating a cognitive representation of the dynamics of road scenes, including anticipating road users' possible choices and actions and flexibly switching attention between different elements of the situation. According to the highly influential Unity and Diversity model of executive functions, executive functions consist of at least three related but dissociable components, namely *switching* of attention or tasks sets, *updating* the contents of working memory, and *inhibition* of task-irrelevant stimuli and habitual responses (Miyake et al., 2000). Here, therefore, we included a measure of each of these aspects of executive functions: Category Switch Task to measure *switching,* Automated Operation Span (AOSPAN) to measure *updating*, and Sustained Attention to Response Task (SART) to measure *inhibition*.

Useful Field of View (UFOV). Like younger drivers, older drivers (60+ years) are another demographic at heightened crash risk (Thompson et al., 2018). Among older drivers, copious evidence suggests that scores on the Useful Field of View (UFOV) task are robustly predictive of both crash risk (for reviews, see Clay et al., 2005; Stefanidis et al., 2023) and hazard perception skill (Anstey et al., 2012; Horswill et al., 2008). However, to our knowledge, the relationship between UFOV and hazard perception in young drivers has not been examined. One study observed a relationship in young drivers between UFOV performance and crashes during a simulated drive (McManus et al., 2015), but it remained unclear whether this relationship was driven by driving skill or other factors, such as driving style. However, there is evidence that cognitive factors predictive of driving behaviour in older drivers are also predictive in younger drivers as well. For instance, Ledger et al. (2019) assessed cognitive predictors of speeding and lane deviation during a simulated drive and found overlap in some of the cognitive predictors of performance for both younger and older drivers (though UFOV was not included). This raises the possibility that UFOV might be a predictor of hazard perception skill in younger drivers, as well as older drivers and hence was included in the present study.

Cognitive Failures. Scores on the Cognitive Failures Questionnaire (CFQ) have been found to predict at-fault workplace safety incidents and car crashes in adults (Larson et al., 1997; Wallace & Vodanovich, 2003). The CFQ was first developed in the 1980s (Broadbent et al., 1982), but there is a revised version of the CFQ with superior psychometric properties that is appropriate for contemporary audiences: The Cognitive Failures Questionnaire 2.0 (CFQ 2.0) (Goodhew & Edwards, 2024). CFQ 2.0 scores better predict objective performance on a non-driving cognitive task than the original CFQ (Goodhew & Edwards, 2024). The present study tested whether CFQ 2.0 scores could predict hazard perception skill – and thus crash risk – in young drivers.

Present study

The goal of the present study was to assess the relationship between select cognitive predictors (measures of switching, updating, and inhibition executive functions, UFOV scores, and CFQ 2.0 scores) and hazard perception skill in young drivers – an aspect of driving skill that is robustly related to crash risk – in a cross-sectional correlational design. The executive function measures were derived from an influential theoretical model (Miyake et al., 2000), while UFOV and CFQ 2.0 were chosen because of their practical utility in predicting crash risk in other populations.

Method

Participants

Participants were required to be between 17-29 years of age, hold a current valid Australian provisional driver’s license, and have normal or corrected to normal vision. Participants were recruited via advertisements placed in Canberra community and university newsletters (e.g., *Canberra Weekly*), Facebook groups (e.g., *Canberra Noticeboard, Canberra Inner North Community group*), websites, and approved physical locations (e.g., noticeboards), and via snowballing.

A power analysis indicated that to have a 95% chance of detecting a medium correlation, N = 134 would be required (G*Power). A medium effect size was chosen because we reasoned that this was the minimum size for the correlation to have practical utility. We then added 10% to the sample to account for data loss and therefore we aimed to recruit 147 participants.

147 participants completed the study. They were aged between 17 and 27 years (*M* = 18.67, *SD* = 1.50), 95 (64.63%) identified as female, 49 (33.33%) as male, 2 as non-binary (1.36%) and 1 as genderfluid (0.68%). In terms of country of birth, 105 participants (71.43%) were born in Australia, with the next most common countries of birth being China (7.48%), India (4.76%), and the United States of America (4.76%). For handedness, 120 (81.63%) reported that they were right-handed, 24 (16.33%) left-handed, and 3 (2.04%) were ambidextrous.

All participants had a provisional Australian driver’s license, of which 59 (40.14%) were Provisional 1 (first year of provisional license after passing the requirements to transition from a learner’s license, e.g., red plates in the Australian Capital Territory, ACT), 85 (57.82%) were Provisional 2 (second and third years of provisional license, e.g., green plates in the ACT), and 3 (2.04%) had Northern Territory (NT) Provisional licenses (which are not differentiated into different categories). In addition, 4 reported a motorcycle license (3 learners) in addition to their car driver's license. Participants reported driving an average of 145.34 kilometres per week over the last year (*SD* = 139.25). Participants were asked how many traffic accidents they had been involved in since obtaining their provisional license (not including accidents involving reversing or parking, or accidents in which their vehicle was not moving), 125 (85.03%) reported 0 crashes, 14 (9.52%) 1 crash, and 6 (4.08%) 2 crashes, 1 (0.68%) 3 crashes, and 1 (0.68%) 4 crashes.

Materials

Hazard Perception Test (Hill et al., 2019). Drivers who can anticipate potentially dangerous situations earlier (e.g., a car turning into one's lane from a side road) have greater opportunity to take evasive action and thus avoid a collision. Consistent with this, the time that it takes a driver to predict hazards in video-based tests is associated with their crash risk (Horswill et al., 2015), and video-based tests of hazard perception skill are used as part of licensing processes in multiple countries. We used the most recent version of the University of Queensland Hazard Perception Test to measure participants' hazard perception skill (Hill et al., 2026; Hill et al., 2019). Hazard perception latencies on this test have been found to differentiate between higher risk (young novice) and lower risk (mid-age experienced) drivers (Hill et al., 2026). Previous versions of the test, developed by the same researchers using the same methodology, but using different footage, have been found to be associated with both the rate of heavy braking events in real-world driving (considered to be an indicator drivers' failure to anticipate dangerous situation sufficiently early (Hill et al., 2019)) and drivers' crash involvement (both retrospectively and prospectively, (Horswill et al., 2015; Horswill et al., 2010)).

In the test, respondents were presented with videos of traffic footage filmed from the perspective of a camera placed in a moving car (akin to the perspective of a driver looking at the road ahead). Participants' task was to click on any road users in the video who were likely to be involved in a traffic conflict with the camera car, which was defined as any situation in which the driver of the camera car would need to take evasive action (e.g., braking) to avoid a collision with the road user. The test consisted of 30 videos[1]. The primary outcome metric was the mean response time to predict traffic conflicts across the videos. Misses were replaced with the mean plus 3SDs, as were response times longer than this (to remove the possibility of slow responders receiving a worse score than non-responders, (as per Horswill & McKenna, 1999)). Response times were converted to z scores and averaged, before being converted back to an overall response time in seconds. The percentage of traffic conflicts detected is also an outcome metric from the test, although this is usually approaching ceiling (by design).

Participants were shown an instruction video prior to starting the test, including two simple examples of traffic conflicts (e.g., a car turning across the path of the camera car). In order to ensure participants had understood the instructions, they had to answer 5 comprehension questions correctly and respond to a practice clip (involving an obvious conflict that any driver who understood the instructions would be expected to respond to). Prior to the Hazard Perception Test, participants also completed a Simple Spatial Reaction Time (SSRT) task that involved them moving the mouse to click on simple stimuli at different locations presented at variable points in time. This provided a control condition that measures their generic baseline motor reaction time using the same response format as the Hazard Perception Test (Poulsen et al., 2010).

Category Switch Task (CST) (Friedman et al., 2008). Here, participants were presented with an object word in the centre of the screen on each trial (e.g., bicycle), and on

---

[1] Due to a technical issue, one of the videos (1 conflict) was missing for participants 1-6. For these six participants, their test score was based on their mean RT to the remaining conflicts.

some trials, their task was to classify the object as *living* or *non-living*, whereas on other trials, it was to classify the object as *bigger* or *smaller* than a basketball. A cue appeared on the screen above the object word at the same time to indicate to participants which classification task was required (heart symbol for living/non-living judgement, an arrow cross symbol for the size relative to a basketball judgement). Participants responded by pressing the 'E' and 'I' keys on the keyboard.

The same 16 target words as Friedman et al. (2008) were used except with the following changes to ensure that the words were appropriate for a contemporary Australian audience: *knob* to *pen*, *alligator* to *crocodile*, *oak* to *tree*, *snowflake* to *raindrop*, and *marble* to *phone. Lizard* was also changed to *butterfly,* because Freidman et al. had *lizard* classified as smaller than a basketball, which is not always the case. Therefore, the 16 target words used here were: *table, bicycle, coat, cloud, pebble, pen, phone, raindrop, shark, lion, tree, crocodile, mushroom, sparrow, goldfish, butterfly.* (In that list, the first four are non-living/larger, next four are non-living/smaller, next four are living/larger, the last four are living/smaller). Prior to the main task, instructions were provided onscreen and participants completed practice blocks (including practicing each task in isolation first).

There were 64 trials in the main test. Judgement type was randomly varied, with the constraint that half of trials were *non-switch* trials, in which the task on that trial was the same as that for the previous trial, and half were *switch* trials, in which the task on that trial was different from the previous trial. Accuracy was designed to be at or close to ceiling in this task, and response time (RT) was the primary dependent variable. Slower responses on switch trials relative to non-switch trial reflected a task-switching cost. The greater the task-switching cost, the lower the task-switching efficiency. Trials with response latencies lower than 100ms were excluded from analysis.

Automated Operation Span (AOSPAN) (Redick et al., 2012). The AOSPAN is a canonical example of a complex span task, where participants are required to juggle performing active mental operations (mental arithmetic) while also encoding and updating information presented sequentially that later needs to be recalled (letters). That is, on each individual trial in the AOSPAN, participants were presented with a simple math equation, such as (8*2) – 3 = ?, and they clicked to continue when they had solved it. Then a number was presented onscreen (e.g., 13), and they indicated whether it was the correct or incorrect answer to the question by clicking 'TRUE' or 'FALSE' labels shown onscreen. Then they were presented with a single letter (to-be remembered). After a variable number of math equation and letter trials, they were presented with all 12 possible letters arranged in a matrix onscreen and instructed to click those that were presented in the order of presentation (recall phase, end of one trial set). The set-size of the trial set refers to the number of individual trials in the set (e.g., set-size of 3 = three letter-number pairs prior to recall phase). The test phase consisted of 75 individual trials arranged as 15 trial sets, comprising 3 repetitions of each of 5 possible set-sizes (set-sizes 3-7), order randomly determined.

Prior to the main block, participants received onscreen task instructions and completed practice trials (including practicing the math and letter components alone and then combined). The outcome measure from the AOSPAN with the best psychometric properties is the total number of letters recalled while maintaining a minimum level of speed and accuracy on the math task (85%, responses within 2.5 standard deviations of their mean RT during practice of math task alone) (Conway et al., 2005; Redick et al., 2012). Higher values indicate greater working memory capacity.

Sustained Attention to Response Task (SART) (Robertson et al., 1997). The SART is a go/no-go task. On each trial a single digit (in varying font size)[2] appeared briefly (250ms) in the centre of the screen, before a mask (a circle with an X) was applied (900ms). Participants' task was to press the spacebar for all digits except 3, for which they were instructed to withhold a response. Each of the 9 digits was presented 25 times, resulting in 225 trials. The key outcome metric is *commission errors*, which is the percentage of digit 3 trials (25 trials) in which the participant erroneously presses the spacebar. Higher commission error rates indicate poorer sustained attention. Prior to the main block, instructions were presented onscreen and participants completed practice trials.

Useful Field of View (UFOV) (Ball et al., 1988; Edwards et al., 2005). The UFOV has multiple variants. Our version consisted of three Subtasks with a fixed peripheral eccentricity. On each trial, the stimulus display was presented briefly before a mask in a staircase design, in which presentation duration was dynamically altered across trials based on participants' performance. Therefore, the key outcome metric was participants' *threshold* (between 17 and 500ms on the 60Hz monitor)*:* the stimulus display time at which they could achieve an accuracy of about 75% correct (lower values = superior performance).

In Subtask 1, participants were presented with a schematic picture of a car or truck in the centre of the screen, and their task was to identify which it was. After stimulus presentation, responses were made by clicking the picture of a car or a truck to indicate their answer. In Subtask 2, participants made the same judgement as Subtask 1 on each trial but also had to identify the location of a single peripherally presented car stimulus. These additional judgements were made by clicking one of eight possible presentation locations (arranged in a circle) on a response screen that appeared after the central car/truck judgement

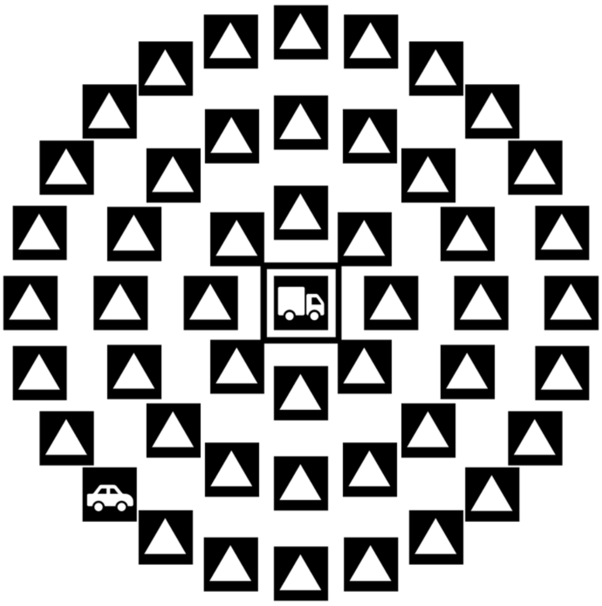
---

[2] The font sizes were percentages of total screen size: 4, 7, 10, 13, 16%. Each font size was equally represented across the test block.

was made. In Subtask 3, participants made the same two judgements as Subtask 2, but now there were additional peripherally presented distractors along with the peripherally presented car (see Figure 1). Previous research shows that, in older adults, Subtasks 2 and 3 are typically predictive of driving-related outcomes (Stefanidis et al., 2023), but in younger drivers only Subtask 3 produces sufficient range for correlational analysis (McManus et al., 2015). Therefore, Subtask 3 was the focus of analysis here. However, the UFOV is typically administered with all Subtasks in sequential order, and thus using one Subtask in isolation could undermine test validity. Hence all three Subtask were administered as blocks always presented in sequential order. Prior to each block, participants received onscreen instructions and completed a practice.

Figure 1

An illustration of the Useful Field of View (UFOV) Subtask 3



Note. This stimulus is presented briefly before a mask. Participants made two responses on each trial: (1) identifying the central stimulus as car or truck, and (2) identifying the location (which one of eight possible locations in the outer ring) that the car was presented at.

Cognitive Failures Questionnaire 2.0 (CFQ) (Goodhew & Edwards, 2024). The CFQ 2.0 is an updated version of the original CFQ (Broadbent et al., 1982), and the updated version has been found to have superior psychometric properties in contemporary samples (Goodhew & Edwards, 2024). It consists of 15 items that reflect failures of attention and memory in everyday life (e.g., *Do you fail to hear people speaking to you when you are doing something else? Do you leave home and later realise you have forgotten something you intended to bring with you?*). Respondents indicated the extent to which they experienced these cognitive failures in the past six months on a five-point scale (*Never, Very Rarely, Occasionally, Quite Often, Very Often).* Total possible scores range from 0-60, where higher scores reflect greater cognitive failures (i.e., poorer attentional control).

Demographic and driving questions. Participants were asked their age, gender, handedness, and country of birth. They reported which type of licence(s) they held, and approximately how long they had spent on each licence type. They reported roughly how many kilometres they had driven per week over the last year, and how many traffic accidents they had been in since obtaining their Provisional licence.

Procedure

Participants were tested individually by an experimenter in a laboratory on the Australian National University (ANU) campus. All participants provided written informed consent prior to participation. Participants completed the tasks in the following order: Hazard perception, UFOV, CST, SART, AOSPAN (via Inquisit's Millisecond player). UFOV, CST, SART, and AOSPAN were either used or adapted from the versions of the tasks available in the Millisecond Test Library as of January 2025 and were run using Inquisit Version 6.6.3. Participants then completed demographic, driving, and CFQ 2.0 questions (via Qualtrics). Participants were provided with AUD $40 cash for their time.

Results

De-identified data are publicly available (https://osf.io/mrfyh/files/osfstorage).

Descriptive statistics. Table 1 shows the descriptive statistics for the measures.

Table 1

Descriptive statistics (N = 147)

Variable	Mean (SD)
Hazard Perception Response Time (s)	7.62 (2.08)
Hazard Perception Hit Rate (%)	99.23 (2.21)
Simple Spatial Reaction Time (s)	0.82 (.10)
CST – Proportion Correct Nonswitch Trials	0.94 (0.05)
CST – Proportion Correct Switch Trials	0.91 (0.07)
CST – Mean RT Nonswitch Trials (ms)	906.32 (193.32)
CST – Mean RT Switch Trials (ms)	1129.12 (277.05)
CST – RT Switch Cost	222.80 (181.63)
AOSPAN Total Letters Correct	57.33 (13.65)
AOSPAN Math Total Errors	4.33 (3.27)
SART Percent Commission Errors	52.05 (21.24)
UFOV Subtask 1 Threshold	16.67 (0)
UFOV Subtask 2 Threshold	17.16 (3.08)
UFOV Subtask 3 Threshold	58.09 (12.41)
CFQ 2.0 Total Score	19.18 (9.72)

Note. CST = Category Switch Task, RT = Response Time, AOSPAN = Automated Operation Span; SART = Sustained Attention to Response Task; UFOV = Useful Field of View; CFQ = Cognitive Failures Questionnaire.

Table 1 shows the Mean and Standard Deviation (SD) of test scores. These indicate that the range of Hazard Perception Hit Rate (close to ceiling) and Hazard Perception Response Times were consistent with what is typically observed for this test (Hill et al., 2019). As is typically observed for the CST, participants' accuracy was high, and responses were faster on Nonswitch trials relative to Switch Trials, indicative of a Switch Cost (Friedman et al., 2008).

The AOSPAN Total Letters Correct is very similar to normative data from over 6000 participants (57.36, SD = 13.65, Redick et al., 2012). While most participants made minimal math errors on the AOSPAN (i.e., trials where they made incorrect math judgement or responded to the math task too slowly), 6 participants did fall below the 85% accuracy criterion (lowest 77%), and so below the correlations below involving AOSPAN total correct letters are considered both with and without these participants.

For the SART, the percentage of commission errors was high (it corresponded to an average of approximately 13 commission errors (out of 25 no-go trials), which is higher than that typically observed in many SART studies (e.g., 4 out of 25 for healthy controls in Robertson et al., 1997, for a review, see Smilek et al. (2010)). The percentage of omission errors (i.e., failing to press the spacebar on go trials) was very low (M = 1.89, SD = 6.92), which is typical in SART designs. Together, this suggests that participants were on average demonstrating a strong tendency to respond on all trials.

For the UFOV, performance in Subtask 1 was at ceiling (all participants achieved the lowest threshold of 16.67ms), and Subtask 2 was near ceiling, consistent with what McManus et al. (2015) observed. This demonstrates that all participants understood and complied with the task instructions. Subtask 3 produced a more substantive range of threshold scores, thus permitting an assessment of the association between performance on this Subtask and hazard perception. Scores on the CFQ 2.0 were similar to that observed in previous work (e.g., *M* = 18.57, *SD* = 8.16 in Goodhew & Edwards, 2024 Study 1).

Associations between Cognitive Predictors and Hazard Perception RT. Calculating bivariate correlations between each cognitive predictor and Hazard Perception Response Time (RT) could inflate Type 1 statistical error rate. Conducting a linear multiple regression with multiple predictors entered simultaneously with an overall model test can help to protect against this. It also allowed us to include Simple Spatial Reaction Time as a covariate to control for individual differences in generic motor response times. This means that any associations between the cognitive predictors of Hazard Perception RT would selectively reflect the time taken to detect hazards.

The criterion was Hazard Perception RT, while the predictors entered simultaneously were: (1) Category Switch Task Mean RT Switch Trials, (2) Category Switch Task Mean RT Nonswitch Trials, (3) AOSPAN total correct letters, (4) UFOV Subtask 3 Threshold, (5) CFQ 2.0 score, and (6) Simple Spatial Reaction Time. We used the Mean RTs in the two different types of trials on the Category Switch Task rather than the subtractive difference score, given that difference scores impair measurement reliability, and this reliability is crucial to the sensitivity of individual difference analyses in finding relationships (Goodhew & Edwards, 2019; Parsons et al., 2019). In this approach, if there is a relationship between Mean RT Switch Trials and Hazard Perception RT while Mean RT Nonswitch Trials are included in the model, then it would indicate a relationship between switching efficiency and hazard perception, uncontaminated by baseline RT differences. We omitted SART Percent Commission Errors from the model given that the descriptive statistics indicate widespread tendencies to press the response key on all trials, contrary to instructions.

This linear multiple regression analysis revealed a significant overall model, *F*(6, 140) = 2.71, *p* = .016, R^2_{adj} = .066. The regression coefficients are shown in Table 2.

Table 2

Regression coefficients (N = 147)

Variable	Standardised regression weight	p-value
CST Mean RT Switch Trials	.15	.238
CST Mean RT Nonswitch Trials	.08	.500
AOSPAN Total Letters Correct	.10	.240
UFOV Subtask 3	-.02	.837
CFQ 2.0	.19	.023*
Simple Spatial Reaction Time	.10	.235

* $p < .05$.

Table 2 shows that only CFQ 2.0 scores were associated with Hazard Perception Response Time, such that those who experienced more frequent cognitive failures in everyday life were slower to detect hazards. If those participants who had 12 or more math errors on the AOSPAN task were excluded, then the results were substantively unchanged for those N = 141 included: There was still a significant overall model, $R^2_{adj} = .06$, $p = .032$, and CFQ 2.0 scores were a significant predictor ($p = .036$), and none of the others were, including AOSPAN Total Letters Correct ($p = .199$).

One possible interpretation of this outcome is that multiple predictors were contributing to explained variance in the criterion, but just via shared rather than unique variance. To test this, CFQ 2.0 was removed as a predictor. The same exclusions for AOSPAN math errors were applied. Now the model became non-significant ($R^2_{adj} = .03$, $p = .095$), and none of the predictors were significant (*p*s>= .196, closest was Simple Spatial RT). This shows that the other cognitive predictors were not operating together to explain variance in Hazard Perception RT. This suggests that CFQ 2.0 scores were uniquely explaining variance in Hazard Perception RT.

Discussion

Hazard perception is a driving skill that is consistently related to crash risk (for a review, see Horswill, 2016). The present study assessed the association between cognitive predictors and measures of hazard perception skill in young provisional drivers who, as a cohort, are at elevated crash risk. These cognitive predictors were selected based on both a theoretical model of executive function, and by identifying measures that have been found to be predictive of crash risk in other populations. Findings indicate that drivers who reported experiencing higher frequencies of cognitive failures in everyday life (e.g., forgetting appointments, forgetting why they went from one part of the house to another) were slower to detect hazards in the video-based hazard perception test. In contrast, UFOV and measures of executive function (including switching and updating working memory) were unrelated to hazard perception skill. The implications of these findings are discussed below.

Cognitive Failures and Hazard Perception. Previous research has shown the subjective experience of cognitive failures in everyday life, as measured by the original CFQ, are associated with important outcomes, such as driver crash risk and at-fault work safety incidents (Wallace & Vodanovich, 2003). However, this is the first study to show that cognitive failures are related to hazard perception (and by extension crash risk) in young drivers.

Further, while the original CFQ has been used extensively for over 40 years, the CFQ 2.0 is a relatively new instrument. The original validation study showed that CFQ 2.0 scores explained more variance in objective performance on a non-driving but safety-relevant task (low prevalence search, akin to that performed by airport baggage screeners searching for weapons) (Goodhew & Edwards, 2024). However, the present study is the first to use this instrument in a road-safety context, and thus the present results provide converging support for the validity of the CFQ 2.0 as assessing individual differences in cognitive function as they relate to meaningful real-world indices, such as hazard perception skill.

Why might cognitive failures relate to hazard perception skill performance, especially when measures of objective performance on cognitive tasks did not? While objective performance on cognitive tasks is thought to gauge an individual's cognitive *ability*, measures of people's subjective experiences of their cognitive function in everyday life, such as that captured by the CFQ, may more strongly reflect their *propensity* to apply their ability (Friedman & Gustavson, 2022). This is a can they/do they distinction: just because a person has strong cognitive abilities, does not always mean that they will apply them, because applying them is effortful. Propensity captures the extent to which they do. From this perspective, the present results could be interpreted as reflecting that cognitive propensity is more predictive of crash risk than cognitive ability in younger drivers. Future research will be required to test this further.

No relationship between Hazard Perception and UFOV or Executive Function. There are two possible interpretations of the absence of these associations. One is that these aspects of cognitive function are truly unrelated to hazard perception skill in young drivers, and another is that they are related to hazard perception skill, but these associations were not observed in the present study for some methodological reason(s). Future research will be required to definitively adjudicate between these two possibilities. However, if we interpret the lack of relationships at face value, then it means that in contrast to UFOV's robust associations with hazard perception in older drivers, UFOV is not associated with hazard perception in younger drivers. This means that UFOV may be selectively sensitive to the cognitive decline that can occur in older adults and impairing their hazard perception.

Further, if we interpret the null relationships as indicative of the absence of a relationship (or at least weaker ones than for cognitive failures), then it bolsters the notion that propensity rather than ability differences are implicated in hazard perception and crash risk in younger drivers. Self-report (questionnaire) measures can be more likely to correlate with other self-report measures and behavioural performance measures with other behavioural performance measures due to shared method variance. The fact that the one relationship that was evident in this study was *across* methods: between a self-report measure and a behavioural performance measure could suggest that this relationship is particularly potent, such that it prevails over any shared method variance.

Limitations and Constraints on Generality. The present study used a convenience sample from a single geographic location and that had particular demographic characteristics (e.g., most born in Australia). Therefore, future research is required to examine the extent to which the present findings generalise into other samples with different characteristics. It should also be noted that the observed association was small-medium in magnitude. This means that the findings should be considered preliminary, and future research is required to assess its robustness.

Conclusions. The present study found that young provisional drivers who experienced greater cognitive failures in everyday life were slower to detect hazards. In contrast, performance on a variety of cognitive tasks were not associated with hazard perception skill. This suggests that young drivers' propensity to apply their cognitive abilities may be especially predictive of crash risk.

Practical recommendations

- These results suggest young provisional drivers would benefit from interventions to improve their propensity to apply their cognitive ability when driving. That is, education programs should be focussed on helping them to understand the importance of applying their full attention to driving.

- Given that the Cognitive Failures Questionnaire 2.0 is a self-report measure of propensity to apply one's cognitive ability, it is unlikely to be useful in any licensing decisions (too easy to "fake good"). However, we are currently working on developing an objective task-performance based measure of propensity to apply one's cognitive ability that could in future be used in this context.
- The present results contrast with our previous project on older drivers (Assessing Predictors of Crash Risk in Older Drivers, funded by the ACT Road Safety Fund, 2022-2023), where cognitive ability measures was associated with crash risk whereas scores on the Cognitive Failures Questionnaire were not. This suggests that the elevated crash risk for these two populations is influenced by different factors. That is, for older adults, it appears that impairments in cognitive ability (rather than propensity) are associated with increased crash risk.
- Given the role of cognitive factors in older drivers' crash risk, we think that there is a need for cognitive ability screening for older drivers to identify those at high crash risk. We plan to submit a project proposal for this year's ACT Road Safety Fund Grant round where the aim is to establish such a fitness-to-drive screening tool for older drivers.

Credit author statement:

S.G: conceptualization, methodology, software, formal analysis, writing – original draft, writing – review and editing, project administration, funding acquisition.

M.E: conceptualization, methodology, writing – review and editing, funding acquisition.

A.H: conceptualization, methodology, software, writing – review and editing, funding acquisition.

M.H. conceptualization, methodology, software, formal analysis, writing – review and editing, funding acquisition.

Acknowledgements: This work was completed with assistance from the ACT Government under the ACT Road Safety Fund Community Grant Program. The authors thank Esther Kong for assistance with the data collection.

References

Anstey, K. J., Horswill, M. S., Wood, J. M., & Hatherly, C. (2012). The role of cognitive and visual abilities as predictors in the Multifactorial Model of Driving Safety. *Accident Analysis & Prevention*, *45*, 766-774. https://doi.org/10.1016/j.aap.2011.10.006

Ball, K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: expanding the useful field of view. *Journal of the Optical Society of America, A, Optics, Image Science & Vision*, *5*(12), 2210-2219.

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, *21*(1), 1-16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Clay, O. J., Wadley, V. G., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, K. K. (2005). Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: current and future implications. *Optometry and Vision Science*, *82*(8), 724-731. https://doi.org/10.1097/01.opx.0000175009.08626.65

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769-786. https://doi.org/10.3758/bf03196772

Di Meco, A., Bennett, J. M., Batchelor, J., Chekaluk, E., Andrews, E., & Habib, J. (2021). The role of cognition for identifying unsafe young drivers. *Safety Science*, *138*, 105099. https://doi.org/10.1016/j.ssci.2020.105099

Edwards, J. D., Vance, D. E., Wadley, V. G., Cissell, G. M., Roenker, D. L., & Ball, K. K. (2005). Reliability and Validity of Useful Field of View Test Scores as Administered by Personal Computer. *J Clin Exp Neuropsychol*, *27*(5), 529-543. https://doi.org/10.1080/13803390490515432

Elander, J., West, R., & French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: an examination method and findings. *Psychological Bulletin*, *113*(2), 279-294. https://doi.org/10.1037/0033-2909.113.2.279

Friedman, N. P., & Gustavson, D. E. (2022). Do rating and task measures of control abilities assess the same thing? *Current Directions in Psychological Science*, *31*(3), 262-271. https://doi.org/10.1177/09637214221091824

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201-225. https://doi.org/10.1037/0096-3445.137.2.201

Goodhew, S. C., & Edwards, M. (2019). Translating experimental paradigms into individual-differences research: Contributions, challenges, and practical recommendations. *Consciousness and Cognition*, *69*, 14-25. https://doi.org/10.1016/j.concog.2019.01.008

Goodhew, S. C., & Edwards, M. (2024). The Cognitive Failures Questionnaire 2.0. *Personality and Individual Differences*, *218*, 112472. https://doi.org/10.1016/j.paid.2023.112472

Hill, A., Horswill, M. S., Salisbury, I., & Scott, T. (2026). Development and evaluation of a new video-based hazard perception test for drivers. *Manuscript in preparation. School of Psychology, The University of Queensland.*

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

Hill, A., Horswill, M. S., Whiting, J., & Watson, M. O. (2023). Assessing speeding propensity via self-report: An on-road validation study of the Driver Behaviour Questionnaire and three speeding-specific measures. *Transportation Research Part F: Traffic Psychology and Behaviour*, *92*, 73-88. https://doi.org/10.1016/j.trf.2022.11.003

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

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

Horswill, M. S., Kemala, C. N., Wetton, M., Scialfa, C. T., & Pachana, N. A. (2010). Improving older drivers' hazard perception ability. *Psychol Aging*, *25*(2), 464-469. https://doi.org/10.1037/a0017306

Horswill, M. S., Marrington, S. A., McCullough, C. M., Wood, J., Pachana, N. A., McWilliam, J., & Raikos, M. K. (2008). The hazard perception ability of older drivers. *The Journals of Gerontology: Series B*, *63*(4), P212-P218. https://doi.org/10.1093/geronb/63.4.P212

Horswill, M. S., & McKenna, F. P. (1999). The effect of interference on dynamic risk-taking judgments. *British Journal of Psychology*, *90*(2), 189-199. https://doi.org/10.1348/000712699161341

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

Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *British Journal of Psychology*, *88*(1), 29-38. https://doi.org/10.1111/j.2044-8295.1997.tb02618.x

Ledger, S., Bennett, J. M., Chekaluk, E., & Batchelor, J. (2019). Cognitive function and driving: Important for young and old alike. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 262-273. https://doi.org/10.1016/j.trf.2018.10.024

McManus, B., Cox, M. K., Vance, D. E., & Stavrinos, D. (2015). Predicting Motor Vehicle Collisions in a Driving Simulator in Young Adults Using the Useful Field of View Assessment. *Traffic Injury Prevention*, *16*(8), 818-823. https://doi.org/10.1080/15389588.2015.1027339

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Advances in Methods and Practices in Psychological Science*, *2*(4), 378-395. https://doi.org/10.1177/2515245919879695

Poulsen, A. A., Horswill, M. S., Wetton, M. A., Hill, A., & Lim, S. M. (2010). A brief office-based hazard perception intervention for drivers with ADHD symptoms. *Australian & New Zealand Journal of Psychiatry*, *44*(6), 528-534. https://doi.org/10.3109/00048671003596048

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*(3), 164-171. https://doi.org/10.1027/1015-5759/a000123

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). "Oops!": Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747-758. https://doi.org/10.1016/S0028-3932(97)00015-8

Smilek, D., Carriere, J. S., & Cheyne, J. A. (2010). Failures of sustained attention in life, lab, and brain: ecological validity of the SART. *Neuropsychologia*, *48*(9), 2564-2570. https://doi.org/10.1016/j.neuropsychologia.2010.05.002

Stefanidis, K. B., Mieran, T., Schiemer, C., Freeman, J., Truelove, V., & Summers, M. J. (2023). Cognitive correlates of reduced driving performance in healthy older adults: A meta-analytic review. *Accident Analysis & Prevention*, *193*, 107337. https://doi.org/10.1016/j.aap.2023.107337

Thompson, J. P., Baldock, M. R. J., & Dutschke, J. K. (2018). Trends in the crash involvement of older drivers in Australia. *Accident Analysis & Prevention*, *117*, 262-269. https://doi.org/10.1016/j.aap.2018.04.027

Wallace, J. C., & Vodanovich, S. J. (2003). Can accidents and industrial mishaps be predicted? Further investigation into the relationship between cognitive failure and reports of accidents. *Journal of Business and Psychology*, *17*(4), 503-514. https://doi.org/10.1023/a:1023452218225

Zicat, E., Bennett, J. M., Chekaluk, E., & Batchelor, J. (2018). Cognitive function and young drivers: The relationship between driving, attitudes, personality and cognition. *Transportation Research Part F: Traffic Psychology and Behaviour*, *55*, 341-352. https://doi.org/10.1016/j.trf.2018.03.013