

# Effects of sleep loss on change detection while driving

## Project Final Report

Prepared for the NRMA-ACT Road Safety Trust

Vanessa Beanland

Ashleigh Filtness

Grégoire Larue

Alana Hawkins

July 2016

## Report Documentation Page

**Date:** 31 July 2016

**Pages:** 81

**Title:** Effects of sleep loss on change detection while driving

**Author(s):** Vanessa Beanland, Ashleigh Filtress, Grégoire Larue, Alana Hawkins

**Sponsoring Organisations(s):** NRMA-ACT Road Safety Trust

### Abstract:

The predominantly urban roads of the ACT create a complex environment in which drivers must quickly detect and respond to changing hazards. This project comprised three experiments designed to assess factors that affect drivers' ability to detect changes in visual information and specifically exploring whether sleepiness impairs change detection, as no previous published research had examined this.

Experiment 1 assessed factors that affect drivers' change detection using photographic stimuli representing urban and rural driving scenes. Accuracy, response time (RT) and eye movements were measured. Participants showed superior change detection in rural compared with urban scenes, and for changes involving road users, animals and traffic lights, compared to inanimate objects (signs and trees).

Experiment 2 used a modified version of the Experiment 1 task to explore the effect of sleep loss on change detection. Participants completed the change detection task twice, once after a normal night's sleep (8 hours) and once following a night of sleep restriction (5 hours). Sleepiness did not impair accuracy, but was associated with increased RT to detect changes in urban scenes. As in Experiment 1, participants were more efficient at detecting changes to other road users than static objects (trees and signs) and were better at detecting changes in rural scenes compared to urban scenes.

Experiment 3 was conducted in the CARRS-Q advanced driving simulator. Participants' ability to detect expected and unexpected changes while driving in simulated urban and rural areas was compared when alert (8 hours sleep) and sleepy (5 hours sleep). Sleep loss did not significantly impair detection of expected changes; however, there was a non-significant reduction in detection of unexpected changes. Participants were better at detecting changes with high safety relevance and in urban areas (where travel speed was low), compared to rural areas (where travel speeds were high).

Overall, this research suggests that drivers are better at detecting changes that involve other road users and targets with high safety relevance. The impact of safety relevance is greatest in demanding situations, e.g. when the visual environment is cluttered or at high travel speeds. There is limited evidence that sleep loss impairs efficiency of change detection in visually cluttered urban scenes. Future research is necessary to understand the vulnerability of visual attention to sleep loss.

### Key words:

change detection; change blindness; sleep restriction; fatigue; driving; hazard perception

### Disclaimer:

This report is disseminated in the interest of information exchange. The views expressed here are those of the authors, and not necessarily those of the Australian National University or the Queensland University of Technology.

## Contents

1. Background .....	6
1.1. Project Objectives .....	6
1.2. Overview of Experimental Series .....	7
1.3. Project Team and Statement of Contributions .....	8
2. Literature Review .....	10
2.1. Driver Sleepiness .....	10
2.2. Change Blindness .....	11
2.2.1. Target Relevance .....	12
2.2.2. Driving Experience .....	13
2.2.3. Familiarity .....	14
2.2.4. Secondary Task Engagement .....	15
2.3. Sleep and Change Blindness .....	15
2.4. Summary and Conclusions .....	16
3. Experiment 1 .....	17
3.1. Background and Rationale .....	17
3.1.1. The Current Study .....	18
3.2. Method .....	19
3.2.1. Participants .....	19
3.2.2. Apparatus .....	19
3.2.3. Stimuli .....	19
3.2.4. Self-Report Measures .....	21
3.2.5. Procedure .....	21
3.2.6. Data Analysis .....	22
3.3. Results .....	22
3.3.1. Participants' Driving Patterns .....	22
3.3.2. Effects of Image Repetition .....	23
3.3.3. Change Detection Accuracy .....	24
3.3.4. Change Detection Response Time (RT) .....	27
3.3.5. Self-Report Measures .....	29

3.3.6. Eye Movements: Fixations on Change Targets .....	29
3.3.7. Eye Movements: Non-Target Fixation Patterns .....	32
3.4. Discussion .....	34
3.4.1. Effects of Driving Environment: Urban vs. Rural .....	34
3.4.2. Effects of Change Safety Relevance .....	34
3.4.3. Effects of Target Type .....	35
3.4.4. Summary .....	36
4. Experiment 2 .....	37
4.1. Background and Rationale .....	37
4.2. Method .....	38
4.2.1. Participants .....	38
4.2.2. Apparatus .....	38
4.2.3. Stimuli .....	38
4.2.4. Self-Report Measures .....	39
4.2.5. Procedure .....	40
4.2.6. Data Analysis .....	40
4.3. Results .....	41
4.3.1. Participants .....	41
4.3.2. Sleep Duration .....	42
4.3.3. Subjective Sleepiness .....	42
4.3.4. Change Detection Accuracy .....	42
4.3.5. Change Detection: Response Time .....	45
4.3.6. Change-absent Trials (Correct Rejections) .....	48
4.3.7. Eye Movements: Fixations on Change Target .....	49
4.3.8. Eye Movement: Non-Target Fixation Patterns .....	52
4.4. Discussion .....	54
4.4.1. Effect of Sleep Loss .....	54
4.4.2. Effect of Target Type .....	56
4.4.3. Effect of Driving Environment: Urban vs Rural .....	57
4.4.4. Summary .....	57

5. Experiment 3 .....	58
5.1. Background and Rationale.....	58
5.2. Method .....	59
5.2.1. Participants .....	59
5.2.2. Apparatus .....	60
5.2.3. Road Network .....	60
5.2.4. Stimuli.....	61
5.2.5. Self-Report Measures .....	63
5.2.6. Procedure.....	64
5.2.7. Data Analysis .....	64
5.3. Results .....	64
5.3.1. Participants .....	64
5.3.2. Sleep Duration .....	65
5.3.3. Subjective Sleepiness.....	65
5.3.4. Change Detection Accuracy .....	65
5.3.5. Unexpected Change Detection Accuracy .....	66
5.4. Discussion .....	67
5.4.1. Effect of Sleep Loss .....	67
5.4.2. Effect of Safety Relevance .....	68
5.4.3. Effect of Driving Environment.....	69
5.4.4. Simulator Methodology Limitations .....	69
5.4.5. Summary .....	70
6. Summary and Conclusions.....	71
7. References .....	75

## 1. Background

The Australian Capital Territory (ACT) *Road Safety Strategy 2011-2020* (ACT Government, 2011) highlights impaired driving as a priority area. Fatigue or sleepiness is one factor that can substantially impair driving performance. Sleepiness-related impairment affects not only motor skills involved in vehicle manoeuvring, but also higher-order cognitive skills including visual attention. Although driver sleepiness is recognised a leading contributing factor in crashes and near-crashes, implicated in approximately 15-30% of all road crashes (Åkerstedt, 2000; Connor, 2009; Horne & Reyner, 1995), it is difficult to objectively measure fatigue and as a result crash records do not accurately reflect the true nature or extent of sleep-related crashes.

Driver sleepiness is particularly likely to affect ACT drivers when making long interstate trips on rural roads, but there is growing recognition within the road safety community that driver sleepiness is also responsible for a high proportion of crashes in urban areas. A survey of ACT and NSW drivers found that approximately 25% of sleep-related crashes occur in residential areas with speed limits of 50 km/h or less, and an additional 30% occur on roads with speed limits between 50 and 80 km/h (Armstrong et al., 2013). Urban roads, such as those that form the majority of the ACT road network, present drivers with unique challenges in that they include a greater variety of road users (i.e., pedestrians, cyclists, motorists) and intersections, have higher visual complexity, and the environment changes more rapidly compared to in rural areas.

In lieu of accurate crash data, experimental evidence is vital for understanding the effects of driver sleepiness and developing targeted interventions. Sleepiness has been associated with significant impairments in simple tasks involving vigilance, psychomotor coordination, and reaction time, as well as more complex cognitive processes such as information processing, memory, and decision making.

One area that has received relatively little attention to date is the effect that sleepiness has on complex visual attention tasks, such as change detection. The ability to detect changes is crucial for safe driving: in order to make safe decisions we must notice when another vehicle has turned onto the road we are driving on, when a bus starts indicating to pull out, or when traffic advisories have been updated with new information. It is difficult to quantify the extent of crashes involving *change blindness* – the failure to detect changes – but research suggests that failure to detect vehicles or hazards is a contributing factor in nearly 10% of serious injury crashes in Australia (Beanland et al., 2013). The current project aimed to address this gap by experimentally quantifying whether, to what extent, and under what conditions sleepiness impairs change detection while driving.

### 1.1. Project Objectives

The broad objective of this project was to examine how sleep loss affects change detection (and, in turn, road safety). The specific objectives were to:

1. Identify which types of visual changes are most difficult to detect in driving scenes.
2. Quantify the extent to which change detection performance varies between driving environments of varying visual complexity (i.e., urban vs. rural roads).

3. Identify the impact of sleep loss on change detection for driving scenes.
4. Assess whether sleep loss has differential effects on performance in different types of driving scenes (e.g., urban vs. rural) or for different categories of stimuli (e.g., vehicles vs. pedestrians).
5. Evaluate drivers' change detection performance during simulated driving, comparing performance while alert and sleepy across both urban and rural driving environments.
6. Provide evidence regarding the effect of sleep loss on drivers' visual attention, which can be used to form policy recommendations, education and awareness campaigns aimed at reducing the incidence of driving while fatigued in the ACT and surrounding regions.

The project objectives were achieved through a series of three experiments, with later studies building on findings from the earlier work. Experiments 1 and 2 were conducted at the Australian National University (ANU) in the Research School of Psychology's eye-tracking lab. Experiment 3 was conducted at Queensland University of Technology (QUT) in the advanced driving simulator at the Centre for Accident Research & Road Safety – Queensland (CARRS-Q).

## 1.2. Overview of Experimental Series

Experiment 1 was designed to assess factors that affect change detection in alert drivers. We used photographic stimuli representing urban and rural driving scenarios, and systematically manipulated the types of changes that occurred in both environments to assess drivers' abilities to detect different types of changing information. The urban vs. rural distinction is important since drivers encounter different types of hazards, and different amounts of visual clutter and complexity, across the two environments. Eye movements were recorded using an Eyelink 1000 eye-tracking system, which provides accurate recording of eye movements in lab-based tasks.

Based on the results of Experiment 1, Experiment 2 used a refined set of stimuli to explore impairments in change detection that result from sleepiness. Experiment 2 used a counterbalanced within-subjects design, so that each driver completed two change detection tasks: one after a good night's sleep and one after experiencing sleep restriction (i.e., a shorter period asleep than normal). The two change detection tasks were matched in terms of the types of changes that occurred and the relative difficulty of detecting these changes. Sleep restriction was achieved by instructing participants to delay their usual bed-time by three hours on the night before the sleep restriction session, but to wake up at their regular time. Compliance with sleep restriction was monitored through use of Body Media SenseWear armbands, which record physical activity, body temperature and galvanic skin response, and therefore provide an objective record of sleep and wake cycles. Eye movements were tracked using the same Eyelink 1000 system as in Experiment 1.

Experiment 3 was conducted in the CARRS-Q advanced driving simulator to explore how sleep loss affects change detection while actually driving. Experiment 3 used a counterbalanced within-subjects design similar to Experiment 2, with all drivers completing three sessions: an initial baseline familiarisation drive and then two experimental drives, one following a good night's sleep and one following a night of sleep restriction. All sessions were at least three days apart, with the order of normal sleep vs. sleep restriction counterbalanced between participants. Each session involved driving

several laps of a route that took them through an urban environment (with roads representing parts of Canberra including civic, the inner north, Parliamentary Triangle, Commonwealth Avenue and Northbourne Avenue) and a rural environment, similar to the types of roads in the region of rural NSW neighbouring Canberra. During the drive a number of expected and unexpected changes occurred and drivers' responses to these changes were recorded. Expected changes occurred following a brief blackout period during the simulation; drivers were required to respond to each blackout by indicating whether a change occurred and describing any changes they observed. Unexpected changes occurred at quasi-random points during the drive and drivers were instructed to indicate and describe any unusual events they noticed. For a subset of participants, eye movements were recorded using the faceLAB eye-tracking system, which enables recording of eye movements during naturalistic tasks.

### 1.3. Project Team and Statement of Contributions

This project was funded through the NRMA-ACT Road Safety Trust's 2014 grants program. Funding for the project was announced in July 2014 and the project commenced shortly thereafter.

The original project team named on the funding application was Dr Vanessa Beanland (ANU) and Dr Ashleigh Filtress (QUT).

Dr Grégoire Larue (QUT) joined the project team in January 2016, when Dr Filtress moved to a new position at Loughborough University in England, and Ms Alana Hawkins (QUT) was employed as a research assistant with substantial responsibility for day-to-day management of Experiment 3, including data collection.

Professor Mike Kyrios (ANU) was appointed as formal project administrator in February 2016, when Dr Beanland moved to a new position at the University of the Sunshine Coast.

Personnel involved in Experiment 1 included:

- Beanland and Filtress were responsible for conceptualisation and design.
- Shannon Webb (ANU research assistant) assisted with the creation of experimental stimuli, specifically taking photographs of driving scenes in the Canberra region.
- Erin Walsh (ANU research assistant) assisted with the creation of experimental stimuli, specifically editing photographs using image editing software to add and insert relevant objects.
- Rhiannon Jeans (ANU research assistant) was responsible for programming the experiment using the SR Builder software, participant recruitment and screening, data collection and initial data processing.
- Jolene Cox (ANU special topics student) recruited a sample of participants to independently rate the safety relevance of the change in each stimulus image used.
- Beanland was responsible for advanced data processing and analysis, with input from Filtress.
- Beanland and Filtress were responsible for write up and presentation of experimental results.

Personnel involved in Experiment 2 included:

- Beanland and Filtress were responsible for conceptualisation and design.



- Jeans was responsible for programming the experiment using the SR Builder software, and for the first half of participant recruitment and screening, data collection and initial data processing.
- Alex Smith (ANU research assistant) was responsible for the second half of participant recruitment and screening, data collection and initial data processing.
- Beanland and Filtness were responsible for advanced data processing and analysis.
- Filtness and Beanland were responsible for write up and presentation of experimental results.

Personnel involved in Experiment 3 included:

- Filtness and Beanland were responsible for conceptualisation and design, with input from Larue.
- Sébastien Demmel, Mindy Li (QUT research associates) and Larue were responsible for programming the driving simulator scenarios.
- Hawkins was responsible for participant recruitment, screening and data collection.
- Wanda Griffin, Oscar Oviedo Trespalacios, David Rodwell and Adrian Wilson were the QUT simulator operators responsible for ensuring the safety of participants during data collection.
- Demmel, Hawkins, Larue and Filtness were responsible for data processing.
- Filtness, Beanland and Larue were responsible for data analysis.
- Filtness, Hawkins, Larue and Beanland were responsible for write up and presentation of experimental results.

## 2. Literature Review

Fatigue or sleepiness is one factor that can substantially impair driving performance. Impairment affects not only motor skills involved in vehicle manoeuvring, but also higher-order cognitive skills including attention. Most research examining the impact of sleepiness on attention has employed basic vigilance paradigms, and there is currently limited research examining the effects of sleepiness and sleep loss on more complex visual attention tasks. The current project sought to address this deficit in the literature by amalgamating domains of driver sleepiness and change blindness, with the aim of exploring how sleep loss affects change detection in complex naturalistic tasks.

### 2.1. Driver Sleepiness

Driver sleepiness represents a significant social and economic cost. Sleep-related crashes account for 15-30% of all crashes (Åkerstedt, 2000; Connor, 2009; Horne & Reyner, 1995) and are associated with higher risk of death and severe injury than other police-reported crashes (Horne & Reyner, 1995). Unlike alcohol intoxication, sleepiness and fatigue cannot be quantified and measured by an index such as Blood Alcohol Concentration (BAC). This is because fatigue can be regarded as a temporary, psychophysiological state that is particularly difficult to quantify in a real-life driving situation (Radun et al., 2013), consequently leading to artificial under-reporting.

Simply analysing crashes based on the ACT's mass databases may be leading to an under-reporting of fatigue-related driving because ACT figures are based solely on police reports, which have been considered to be an under-estimate of the true number of fatigue-related crashes (Attewell et al., 2001). Given the difficulties in objectively measuring fatigue and sleepiness in the real world (as compared with, for example, speed or intoxication where police may monitor and objectively record the extent of violation), education and awareness campaigns that encourage drivers to monitor their own fatigue levels are likely to be the most efficient way of reducing driver fatigue in the community.

A recent survey of 1,609 drivers from ACT and NSW found that most had experienced sleepiness while driving in the past 5 years (Armstrong et al., 2011). ACT drivers were more likely than NSW drivers to experience driver sleepiness (71% vs. 62%). More concerning, ACT drivers were also more likely to continue driving despite feeling sleepy and were more likely to report multiple sleep-related "close calls" (i.e., near-crash incidents). Overall these results suggest that ACT drivers do not take adequate precautions to avoid driving while fatigued and that more work is needed to raise driver awareness of the negative consequences of driving while sleepy.

The under-reporting of fatigue in police data means that experimental research is vital for understanding the road safety implications of driver sleepiness. For example, experimental research has revealed that vigilance decrements after 17 hours awake is equivalent to that of a driver with 0.05% BAC (Dawson & Reid, 1997). This suggests that Canberra drivers who plan to drive to the NSW south coast on Friday evening after a full day's work may experience impairment equivalent to illegal levels of intoxication.

There are obvious safety concerns in conducting driver sleepiness research on real roads. Driving simulators are a safe alternative, as they permit researchers to create controlled environments in which

they can measure the impact of sleepiness on driving (Liu et al., 2009). Driving simulator studies reveal that vehicle control is impaired following sleep loss (Anund et al., 2008), resulting in increased lane deviations (Filtner et al., 2012). Most previous research investigating driver sleepiness focused on rural highway driving, as sleep-related crashes are particularly likely to occur during monotonous driving conditions. However, among ACT and NSW drivers approximately 25% of sleep-related crashes occur in residential areas with speed limits of 50 km/h or less and a further 30% occur on roads with speed limits between 50 and 80 km/h (Armstrong et al., 2013). As such, driver sleepiness in urban environments has a significant impact on overall road safety, but is an area that has been neglected by previous research and policy. Consequently there is poor understanding of how sleepiness and fatigue affect driver performance in urban environments and there are no targeted countermeasures aimed at reducing driving while fatigued in urban environments.

Although it has been established that sleepiness impairs vehicle control (i.e., manual handling and manipulation of controls) there are also implications for other vital skills necessary for safe driving. It is well established that performance on simple vigilance and reaction time tasks is impaired by sleep loss (e.g., Belenky et al., 2003; Dinges et al., 1997; Van Dongen et al., 2003). More complex cognitive processes, such as information processing and planning ability, are also impaired following sleep loss (Horne, 2012). Finally, sleep loss also impairs one's ability to complete dual task paradigms (Haavisto et al., 2010) and makes drivers more susceptible to distraction, leading them to make a greater number of glances away from the road (Anderson & Horne, 2013).

Furthermore, sleep loss has implications for vision and oculomotor control (i.e., eye movements and blinks). Sleep loss increases double vision (Clark & Warren, 1939) and exoropia, or divergence of the eyes outward (Horne, 1975). Recently it has been noted that sleep deprivation leads to decreased oculomotor function (De Gennaro et al., 2000; Fransson et al., 2008), which impairs visual search performance (De Gennaro et al., 2001). This has prompted the suggestion that oculomotor control could be used as a fatigue detection measure (Goldich et al., 2010; McClelland et al., 2010). The interaction between sleep loss and eye movements suggests a mechanism by which sleep loss could influence change blindness, since change blindness is also significantly influenced by eye movements, as discussed further in Section 2.2. This has potential implications for road safety as saccadic velocity (i.e., speed of eye movements) is negatively correlated with simulator vehicle crashes (Rowland et al., 2005; Russo et al., 1999, 2003). Although sleep loss impairs several skills that are vital to safe driving in urban environments, no previous research has shown a direct relationship between sleepiness and urban driving safety.

## 2.2. Change Blindness

Change blindness is psychological phenomenon in which observers either completely fail to detect changes within a visual scene, or experience a substantial delay in detecting a change within their visual environment (Rensink et al., 1997). Change blindness is particularly likely to occur when visual changes take place during a disruption to the visual scene, such as when a person is blinking, making an eye movement, or has their view obscured briefly (e.g., McConkie & Currie, 1996; Pashler, 1988; Rensink et

al., 1997; Simons & Levin, 1998), as the disruption masks visual transients that would otherwise make the change obvious to the observer.

Change blindness can occur for both expected and unexpected changes, across a wide range of visual stimuli including simple arrays of letters and digits (Pashler, 1988), photographs (Rensink et al., 1997) and even a person in a real-life conversation (Simons & Levin, 1998).

Several previous studies have examined the incidence of change blindness while driving using a range of methods, including both driving simulation and computer-based experiments similar to the methods employed in the current project. The most common methods used in driving-related change detection research are flicker tasks, one-shot tasks, and simulated driving scenarios.

In *flicker tasks*, two alternating images are presented for a fraction of a second each (typically 240-500 ms), separated by a brief (80-500 ms) blank screen that serves to mask visual transients (Rensink et al., 1997). The sequence “flickers” between the two images until the observer determines whether the two images are the same or different.

*One-shot tasks* use a similar format, with two images presented for a fixed duration separated by a blank screen, but each image is presented only once and stimulus durations are often longer (e.g. 10-15 s; Zhao et al., 2014). As there is limited opportunity to compare the images, accuracy is typically lower in one-shot tasks compared with flicker tasks.

*Simulated driving paradigms* embed change detection tasks within a driving simulator scenario. Some simulator studies mask changes with brief occlusion periods (Lee et al., 2007; Shinoda et al., 2001; Velichkovsky et al., 2002; White & Caird, 2010), similar to the blank screens used in flicker and one-shot tasks, whereas others have changes occur more naturalistically, for example changing a sign between repeated drives on the same road (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007) or during an eye movement (Velichkovsky et al., 2002).

Previous research has examined how change detection in driving scenes is affected by several variables, including target relevance, driving experience, familiarity with the road environment, and secondary task engagement. Key findings pertaining to each of these topics are summarised in the following subsections.

### **2.2.1. Target Relevance**

A robust finding in change blindness research is that observers are faster and more accurate at detecting changes to targets that have greater relevance, such as targets that are central to understanding the scene (Rensink et al., 1997) or targets that are personally meaningful (Marchetti et al., 2006). Similarly, studies consistently reveal that observers are faster and more accurate at detecting changes in road scenes when the targets are driving-relevant, compared with driving-irrelevant targets (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002; Zhao et al., 2014). One caveat to these findings is that many studies use quite broad definitions of “relevant” and “irrelevant”. Examples of relevant targets include vehicles, pedestrians and road signs, whereas examples of irrelevant targets include buildings, dumpsters and mailboxes (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002). This raises a potential confound, in that the irrelevant targets are all stationary objects, which

are typically positioned away from the road and, in turn, farther from the driver's central focus. Moreover, these studies group together several types of driving-relevant targets, which vary considerably in their potential relevance to driving safety.

Velichkovsky et al. (2002) also found faster and more accurate change detection for task-relevant stimuli in road scenes, and noted that change blindness was stronger in dynamic stimuli compared to static stimuli. This suggests that change blindness may be more likely to occur during simulated driving (and potentially real driving) than in lab-based experiments. This is, in part, due to eye movements: change blindness can occur as a result of saccadic suppression (i.e., when visual input from the retina to the brain is temporarily suppressed during saccades or eye movements). For this reason, it is particularly relevant to explore the role that eye movements play in successful vs. unsuccessful change detection.

Two simulator studies have provided more systematic manipulation of safety-relevance within a single class of targets (Lee et al., 2007; Shinoda et al., 2001).

In the first study, by Shinoda et al. (2001), the same change occurred during each trial – a “no parking” sign changed into a “stop” sign – but target placement was systematically manipulated to alter its relevance in relation to drivers' expectations. Drivers were significantly less likely to notice the changing sign when they were following another car, or when it occurred mid-block, compared with when it occurred at an intersection (Shinoda et al., 2001). Arguably, stop signs are equally relevant regardless of where they appear; however, drivers *expect* signs at intersections to convey more meaningful information (e.g., whether one has priority or must give way to other traffic).

In a later study, Lee et al. (2007) tested drivers' ability to detect changes to vehicles, which were either parked, moving ahead of the participant, or moving behind the participant. Drivers were most sensitive to lead vehicles moving closer to them (i.e., simulating a sudden braking movement) and were least sensitive to changes involving parked vehicles. This suggests that drivers are more efficient at detecting safety-relevant changes; however, the authors noted that target location co-varied with safety relevance, and as such the results cannot be solely attributed to safety relevance without further research (Lee et al., 2007).

Although several studies have compared change blindness for task-relevant vs. irrelevant stimuli, the choice of stimuli has mainly been restricted to objects that have indirect relevance to road safety, such as road signs. As such, one aim of the current project was to more comprehensively explore how the nature of the stimuli affects change blindness vs. change detection, by comparing stimuli with varying levels of relevance to the task of safe driving.

### **2.2.2. Driving Experience**

Change blindness research in non-driving domains consistently indicates that domain-experts are less susceptible to change blindness compared to domain-novices, but only for expertise-related changes (Feil & Mestre, 2010; Reingold et al., 2001; Werner & Thies, 2000). For instance, American football experts are faster than non-experts at detecting changes to football-related images that meaningfully alter game formations, but not at non-meaningful or non-football-related changes (Werner & Thies, 2000). Comparable findings have been obtained for chess masters (Reingold et al., 2001) and advanced physics students (Feil & Mestre, 2010). Based on this it seems logical that driving experience would

similarly influence change detection ability in driving-related scenes; however, empirical findings have been mixed (Zhao et al., 2014).

One method of examining effects of driving experience is to compare drivers with non-drivers, that is, people who have never held a driver's licence. An English study comparing non-drivers and drivers found no significant association between driving experience and performance on a driving-related flicker change detection task, although both groups were faster at detecting driving-relevant compared with irrelevant changes (Galpin et al., 2009). The authors suggested that their driver group may not have had sufficient experience (average 70 months) to demonstrate superior performance. Following this, a Chinese study compared change detection ability in non-drivers and drivers with an average of 33 months experience (Zhao et al., 2014). The Chinese study used a one-shot task and inserted a central fixation point on half the trials. Drivers and non-drivers performed similarly on trials with no fixation point, replicating Galpin et al.'s (2009) results. When the fixation point was present, drivers and non-drivers also performed similarly for centrally-located and driving-irrelevant changes, whereas non-drivers were significantly less accurate than drivers at detecting driving-related and peripheral changes (Zhao et al., 2014). The authors suggested driving experience helps facilitate more efficient processing of driving-related and peripheral elements while fixating centrally.

Beyond comparing drivers and non-drivers, another method for studying experience effects is to compare change detection abilities between drivers with varying experience levels. In a US study comparing young novice drivers (average 6 months experience) to more experienced young drivers (average 7 years' experience), both groups performed similarly on driving-related changes but novice drivers were less accurate at detecting irrelevant changes (Mueller & Trick, 2013). One explanation is that experienced drivers are more efficient at processing driving-related information, which means they have greater cognitive capacity remaining for processing irrelevant information. This is consistent with Lavie's (1995) load theory, which posits that task-irrelevant information will only be selected into conscious awareness under conditions of low task load (i.e., when the primary task is less cognitively demanding). It is also consistent with Zhao et al.'s (2014) findings, whereby drivers showed superior detection of peripheral changes compared with non-drivers.

Finally, an Australian study found that after accounting for simple reaction time differences, drivers with less than 3 years' experience were significantly *faster* at detecting driving-related changes, compared with drivers who had more than 10 years' experience (Wetton et al., 2010). However, it is worth noting that this study's "novice" group actually had a similar level of experience to the participants considered experienced drivers in other studies (e.g., Zhao et al., 2014) and were on average 19 years younger than the comparison group of experienced drivers in the same study (Wetton et al., 2010). Overall it seems that differences in change detection ability may be most likely to emerge when comparing drivers to either non-drivers, or those with only a few months' experience.

### **2.2.3. Familiarity**

A few studies have explored the effect of environmental familiarity on change detection while driving (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007). These studies use broadly similar methodology: all recruited groups of drivers to complete 20-25 simulated drives over a period of



several days or weeks. Whereas most studies assess short-term changes – i.e., detecting that an object has appeared, disappeared, moved or changed within the past second – studies that explore the effects of familiarity usually test long-term change detection, such as whether drivers notice that a speed limit has been altered since the previous time they drove on that road.

Overall, these studies suggest that repeatedly driving the same route increases drivers' ability to recognise certain aspects of the environment but impairs others. For instance, drivers are better at recognising which roads signs belong on a route (Martens & Fox, 2007) and are faster at detecting a target vehicle when they are more familiar with the route (Charlton & Starkey, 2013). However, these benefits appear to be offset by substantial change blindness to other aspects of the environment, particularly road signs, even when the changes has clear safety relevance. For instance, many drivers failed to detect when an intersection sign changed from granting them priority to requiring them to give way (Martens & Fox, 2007), when speed limits on dynamic speed signs changed (Harms & Brookhuis, 2016), or when the sign's language changed from English to German (Charlton & Starkey, 2013). Drivers also exhibited robust change blindness to the addition or removal of roadside buildings, but were much better at detecting changes to road markings, even after repeated exposure (Charlton & Starkey, 2013). This suggests that when driving on familiar routes, drivers pay relatively less attention to the roadside – including safety-relevant signs – but maintain focus on the road itself.

#### **2.2.4. Secondary Task Engagement**

Studies examining the impact of secondary task engagement on driving-related change detection have all indicated that engagement in a cognitively demanding secondary task significantly impairs change detection (e.g., Lee et al., 2007; McCarley et al., 2004; Richard et al., 2002; White & Caird, 2010). These effects have been demonstrated using flicker tasks with photographs depicting road scenes (McCarley et al., 2004; Richard et al., 2002) and also in driving simulator scenarios in which changes occur after brief blackouts (100ms to 1s; Lee et al., 2007; White & Caird, 2010).

The specific aspects of change detection affected by dual-task engagement differ between studies, and include accuracy, sensitivity and response time. Early research on this topic found that concurrent engagement in an auditory working memory task resulted in slower change detection but did not affect accuracy (Richard et al., 2002). However, subsequent research has found that responding to auditory messages and engaging in hands-free phone conversations (but not passively listening to a conversation) impairs change detection accuracy (Lee et al., 2007; McCarley et al., 2004). Notably, drivers were equally likely to fixate change targets when talking on a phone, but failed to consciously process the change (McCarley et al., 2004). Finally, White and Caird (2010) found that young adult drivers who were accompanied by an attractive opposite-sex passenger were less likely to detect hazards, compared to participants who were driving alone. Together these findings suggest that driver distraction can increase change blindness and “looked-but-failed-to-see” errors.

### **2.3. Sleep and Change Blindness**

There has been almost no previous research examining the relationship between sleep and change blindness or related phenomena.

Some research has compared change detection performance between good sleepers and people with insomnia or other sleep disorders (e.g., Marchetti et al., 2006). This research found that insomniacs are better at detecting changes to sleep-related stimuli, which suggests that they have an attentional bias towards sleep-related stimuli. However, this study did not look at change detection more broadly so it is unclear whether sleep loss affects generic change detection abilities. Further, the study looked specifically at people with sleep related disorders, not at good sleepers who have experienced a temporary period of sleep loss, so the results are not generalizable to the broader population.

## 2.4. Summary and Conclusions

In-depth research has shown that driver fatigue due to sleepiness or sleep loss is a leading contributory factor in road crashes and should be considered a significant road safety issue for all road users. While archetypal sleep-related crashes (i.e., where a single vehicle runs off the road into a tree on a monotonous country road) are relatively easy to identify and are well researched, atypical sleep-related crashes on urban roads are poorly understood and hard to identify.

This project was designed to a significant gap in the literature by improving understanding of visual attention impairments relating to sleep loss in both urban and rural driving environments. Given the relationship between change blindness and eye movements, it was hypothesised that sleep loss would impair change detection, resulting in reduced accuracy and/or longer response times in change detection tasks. Individuals experiencing sleep loss typically demonstrate increased blink rate and spend a longer percentage of time with their eyes closed. These changes in blink patterns may result in slower change detection and increased change blindness, since change blindness is more likely occur during blinks and eye movements.

In addition to changes in blink patterns, it is possible that sleepy drivers may show differential patterns of visual scanning. In particular, they may attempt to compensate for their sleepiness by focusing their eyes on the road ahead, at the expense of detecting peripheral information on the roadside. This type of compensatory behaviour is observed during distracted driving (e.g., Engström et al., 2005) and distracted drivers show greater impairment at responding to peripheral vs. central hazards (Haque & Washington, 2013). If similar results are observed as a result of sleep loss, then it is likely that drivers would have greater impairment at detecting changes to peripheral objects including signs and hazards on the roadside.



## 3. Experiment 1

### 3.1. Background and Rationale

The ability to detect changes is crucial for safe driving. In order to make appropriate decisions we must notice when another vehicle pulls out ahead of us, when an in-vehicle alert appears, or when advisory signs have been updated. Research examining change detection while driving (e.g., Charlton & Starkey, 2013; Galpin et al., 2009; Velichkovsky et al., 2002; Zhao et al., 2014) suggests that drivers often experience *change blindness*, which is delayed or failed change detection (Rensink et al., 1997). Although it is difficult to quantify the extent of crashes involving change blindness, accurate change detection is associated with safe decision-making (Caird et al., 2005; Edwards et al., 2008) and in-depth crash analyses suggest approximately 9% of serious injury crashes involve a driver failing to detect hazards (Beanland et al., 2013).

Several paradigms have been used to explore change blindness (for a review see Jensen et al., 2011). The diversity of paradigms stems from the fact that change blindness can occur for expected or unexpected changes, and can result from various visual disruptions including blinks, saccades, or occlusion (Beanland et al., 2015).

The most common research methods used in driving-related change detection research are flicker tasks, one-shot tasks, and simulated driving scenarios. In *flicker tasks*, two alternating images are presented for a fraction of a second each (typically 240-500 ms), separated by a brief (80-500 ms) blank screen that serves to mask visual transients (Rensink et al., 1997). The sequence “flickers” between the two images until the observer determines whether the two images are the same or different. *One-shot tasks* use a similar format, with two images presented for a fixed duration separated by a blank screen, but each image is presented only once and stimulus durations are often longer (e.g. 10-15 s; Zhao et al., 2014). As there is limited opportunity to compare the images, accuracy is typically lower in one-shot tasks than in flicker tasks. *Simulated driving paradigms* embed change detection tasks within a driving simulator scenario. Some simulator studies mask changes with brief occlusion periods (Lee et al., 2007; Shinoda et al., 2001; Velichkovsky et al., 2002; White & Caird, 2010), similar to the blank screens used in flicker and one-shot tasks, whereas others have changes occur more naturalistically; for example, changing a sign between repeated drives on the same road (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007) or during an eye movement (Velichkovsky et al., 2002).

Previous research has examined how change detection in driving scenes is affected by several variables, including target relevance, driving experience, familiarity with the road environment, and secondary task engagement. Section 2.2 provides a full review of the relevant literature, but key findings are summarised briefly below.

Research has consistently demonstrated that observers are faster and more accurate at detecting changes that have greater relevance to driving (Galpin et al., 2009; Lee et al., 2007; Mueller & Trick, 2013; Shinoda et al., 2001; Velichkovsky et al., 2002; Zhao et al., 2014). However, these studies typically employ targets with only indirect relevance to driving, often have systematic differences between relevant and irrelevant targets, and collapse results across several distinct types of targets to form their

“relevant” and “irrelevant” categories. This highlights several avenues for conducting more nuanced investigation into the relationship between change detection and target relevance.

Findings regarding the effects of driving experience on change detection are mixed (Zhao et al., 2014), in part due to the fact that different studies use varying methods and forms of comparison (e.g., some compare drivers and non-drivers, others compare drivers of varying experience levels). Some studies find that driving experience is associated with superior change detection, but only for certain types of changes (Mueller & Trick, 2013) or under specific conditions, such as in the presence of a central fixation point (Zhao et al., 2014), whereas other studies find no relationship between driving experience and change detection (Galpin et al., 2009; Wetton et al., 2010). Given these inconsistencies, the most sensible approach for research exploring driving-related change detection is to exclude novice drivers and those with very little experience, to ensure that there is no potential for experience-related effects to confound the study’s results.

### **3.1.1. The Current Study**

Based on the review of previous change detection research in Section 2, it is apparent that change blindness occurs in driving environments, but that the extent of change blindness varies depending on characteristics of the changed object. Characteristics such as object size or physical salience (Koustanai et al., 2012) do not predict the efficiency of change detection in naturalistic tasks. Rather, semantic object properties (e.g., relevance to driving) influence the likelihood and speed of change detection. Previous studies examining this have either defined task relevance quite broadly (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002; Zhao et al., 2014) or have used only a single class of targets (Lee et al., 2007; Shinoda et al., 2001), so there is scope for more systematic investigation of the relationship between target characteristics and change detection.

The current study was designed to assess drivers’ change detection efficiency in urban and rural driving scenes across a range of target types including vehicles, vulnerable road users, signs, and roadside objects. All of these targets are potentially relevant to safe driving depending on the context in which they appear, so we systematically manipulated the change context within each category of targets. This resulted in a total of seven target categories (cars, motorcycles, road signs, traffic lights, pedestrians, animals, and trees), with half of the trials in each category containing changes that have high potential for safety impact (i.e., requiring monitoring of a potential hazard or a response by the driver) and half containing changes that have low or no potential for safety impact (i.e., the driver can continue without any change in behaviour or situation awareness). This allowed us to explore which factor is more influential in change detection, the type of target or its potential safety impact, and whether these two factors interact. In addition to standard measures of accuracy and response time (RT), eye movements were recorded to provide a more comprehensive understanding of change detection occurs (i.e., by examining “looked-but-failed-to-see” errors and implicit capture of attention).

## 3.2. Method

### 3.2.1. Participants

Twenty-six drivers (15 female, 11 male) aged 20-43 years ( $M = 22.9$ ,  $SD = 4.7$ ) provided informed consent and participated voluntarily in exchange for AUD\$20. Data from one additional participant was discarded due to technical errors. All participants had normal or corrected-to-normal visual acuity (as measured using a near vision chart), held a current full unrestricted Australian driver's licence, and drove at least once a week within the Canberra region. Ethical aspects of the research were approved by the Australian National University Human Research Ethics Committee (protocol 2014/458).

### 3.2.2. Apparatus

Visual stimuli were presented on a 27" Apple iMac desktop computer. An Eyelink 1000 eye-tracker, with a reported spatial accuracy within  $0.25$ - $0.5^\circ$ , was used to monitor eye movements at a temporal frequency of 1000Hz. Head position was fixed using a chinrest with a viewing distance of 95cm, yielding a display area of  $30.3^\circ \times 19.4^\circ$  visual angle. Stimulus presentation and data acquisition were controlled via SR Research Experiment Builder.

### 3.2.3. Stimuli

Experimental stimuli included 200 image pairs depicting driving scenes, which constituted 50 urban change-present pairs, 50 rural change-present pairs, 50 urban change-absent pairs and 50 rural change-absent pairs. Each image subtended  $23.0^\circ \times 17.5^\circ$  visual angle and was taken using a digital camera mounted on the dashboard of a station wagon. Urban images were taken in central Canberra (civic, inner north, Parliamentary Triangle) and rural images were taken on rural roads in surrounding regions. In change-absent pairs the two images displayed were identical, whereas in change-present pairs one of the images was edited to add, remove or alter a single driving-relevant target. Images used were selected from a larger sample ( $N > 2000$ ) of photographs so that change-absent and change-present images could be matched in terms of the roads, road users and visual complexity within scenes.

Within both the urban and rural environments, five types of target objects were changed. In the urban scenes change targets were either cars, motorcycles, road signs, traffic lights or pedestrians, with 10 images for each category. In the rural scenes change targets were either cars, motorcycles, road signs, trees or animals, again with 10 images for each category. For the three categories that occurred in both urban and rural scenes (i.e., cars, motorcycles, and road signs) the nature of the changes was matched so that equivalent changes occurred in both environments.

Within each target type, the potential safety impact of the change was systematically manipulated, so that half the images contained a change with high safety impact (e.g., vehicle appears/disappears immediately in front of the participant, change to speed limit sign) and half contained a change with low potential safety impact (e.g., parked vehicle appears/disappears, change to bicycle lane advisory sign content). The key differentiator between high- and low-impact images was that high-impact changes would require a driver to change their behaviour (e.g., adjust travel speed, brake, monitor a potential hazard), whereas low-impact changes did not require any changes to behaviour or situation awareness. Because previous research has suggested that drivers' subjective risk perceptions do not always align

with objective assessments of risk (Charlton et al., 2014), we recruited a separate sample of 21 experienced drivers aged 25-40 years to rate the safety relevance of each change on an 11-point scale from 0 (not at all safety relevant) to 10 (highly safety relevant). The average rating for each image pair was then used as the safety relevance rating for each trial.

Image pairs were presented using a “flicker” sequence, in which one image was presented for 500ms, followed by a 500ms blank grey screen, followed by the second image for 500ms and then another 500ms blank (see Figure 3-1). The cycle of alternating images and blanks continued until the participant responded, or for 30 s, whichever occurred first. Participants were instructed to decide as quickly as possible whether a change occurred and then immediately press the space bar to register their decision. They were then prompted to report whether a change occurred and, if applicable, the change target. If participants failed to respond within 30 s the program automatically proceeded to a response screen that asked them to indicate whether a change occurred. Available response options included “yes” and “no” for whether a change occurred, and “vehicle”, “motorcycle”, “bicycle”, “person”, “animal”, “tree”, “building”, “sign”, and “traffic light” for change target. Change-present trials were considered “correct” if the observer correctly identified the change target, but were considered “incorrect” if they reported no change or failed to select the correct change target. Change-absent trials were considered “correct” if the observer reported no change, and were considered “incorrect” if they indicated a change occurred (this form of error was rare, occurring on 0.7% of trials).



*Figure 3-1.* Example trial sequence from Experiment 1, showing an urban scene where a change occurs between image A and image B (the blue car appears/disappears).

The experiment contained 220 trials, which comprised 200 trials with unique image pairs (100 change-present, 100 change-absent, as described above) and 20 trials with repeated images (10 change-present, 10 change-absent). Unique and repeated images were analysed separately.

Repeated images were included because the proposal for Experiment 2 required participants to complete two change detection sessions. Performance for the first vs. second presentation of repeated images was therefore compared to assess whether it is feasible for participants to complete two change

detection sessions using identical stimuli, or whether it would be necessary to develop two separate but equivalent stimulus sets for use in each session.

The experimental task was preceded by 5 practice trials (3 change-present, 2 change-absent), which used novel images taken from a previous unrelated change detection study.

#### **3.2.4. Self-Report Measures**

Participants completed a brief demographic questionnaire and two self-report inventories, the Driver Behaviour Questionnaire (DBQ; Lajunen et al., 2004; Lawton et al., 1997; Mattsson, 2012; Parker et al., 1995) and the Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982).

The DBQ requires respondents to rate their frequency of engaging in 28 aberrant driving behaviours on a 6-point Likert scale from 0 (never) to 5 (nearly all the time). Previous research has typically found that in English-speaking populations this scale reveals four subtypes of aberrant driving behaviour (Beanland et al., 2014b): *Ordinary Violations*, or deliberately disregarding road rules and norms; *Aggressive Violations*, involving hostility towards other road users; *Errors*, which are dangerous non-deliberate acts, such as failing to search for or detect oncoming traffic before entering an intersection; and *Lapses*, which are relatively minor failures, such as misreading road signs or forgetting where one's car is parked. For the current study, the Errors and Lapses subscales were of particular interest.

The CFQ requires respondents to rate their frequency of 25 lapses of attention, perception and memory in everyday life on a 5-point Likert scale from 0 (never) to 4 (very often). Originally it was claimed that the scale measured a unitary construct, with specific subfactors varying between populations (Broadbent et al., 1982). Subsequent studies have found that multi-factor solutions fit the data better than single-factor solutions (Bridger et al., 2013; Wallace, 2004); however, the specific factor structure varies between populations and even within populations over time (Bridger et al., 2013). Given this inconsistency, and the fact that overall CFQ scores have been found to significantly predict performance in some visual attention tasks (e.g., Forster & Lavie, 2007), for the current study overall CFQ scores were analysed.

#### **3.2.5. Procedure**

Participants were tested individually in a quiet laboratory, which was completely dark during the eye-tracking experiment. After providing informed consent and completing the visual acuity screening, participants completed the self-report measures (i.e., demographic questionnaire, DBQ, CFQ).

After completing the questionnaires participants were seated in front of the computer with their head position stabilised using a chinrest. The eye-tracker was individually calibrated for each participant using a 16-point calibration grid and then validated to ensure that average gaze error was  $<0.5^\circ$ , which is within the margin of acceptable error specified by the manufacturer. Each trial commenced with a drift check to ensure gaze calibration accuracy was maintained and the system was manually recalibrated if the error exceeded  $1.0^\circ$  for three consecutive trials. Participants then completed the experiment, with breaks offered every 55 trials.

### 3.2.6. Data Analysis

Accuracy, response time (RT) and eye movements to the change target were analysed using Generalized Estimating Equations (GEE; Liang & Zeger, 1986), an extension of the general linear model that permits analysis of repeated measurements where not all participants contribute the same number of observations (i.e., trials) to the dataset. Binary logistic GEE functions similarly to binary logistic regression, but because GEE permits repeated measurements it can be used to assess whether the probability of a binary outcome (e.g., change detection, fixating a change target) differs according to within-subjects variables (e.g., target type). Linear GEE functions similarly to repeated-measures analysis of variance (RM-ANOVA) and can be used to assess whether continuous variables (e.g., RT, dwell time on target) differ according to within-subjects variables. The crucial difference between GEE and ANOVA is that GEE is based on individual trials, whereas ANOVA is based on averages and requires that all participants have data in each condition (otherwise all of their data is excluded from the analysis). This is problematic for change detection paradigms as RT analyses include only correct trials, but some observers fail to detect all targets of a specific type (in the current study, this was common for the “tree” changes). GEE is therefore useful as it can accommodate missing data ranging from single trials to entire conditions, and provides greater statistical power compared with RM-ANOVA (Ma et al., 2012).

Correlations and paired *t*-tests were used for other measures where overall performance was of interest (e.g., correlations between cognitive failures and change detection performance). All analyses were conducted in IBM SPSS Statistics 22. An alpha level of .05 was used to assess statistical significance.

## 3.3. Results

### 3.3.1. Participants' Driving Patterns

Participants had an average self-reported driving frequency of 4.9 hours (*SD* = 3.3; range 1-18 hours) or 182 km (*SD* = 133; range 20-500 km) per week. As shown in Figure 3-2, participants drove most frequently on urban roads. Nearly 90% reported that they drove on urban 60 km/h roads frequently or all the time, and 58-65% reported driving on higher speed urban roads frequently or all the time. In contrast, over 90% reported that they drove on rural roads occasionally, hardly ever, or never.



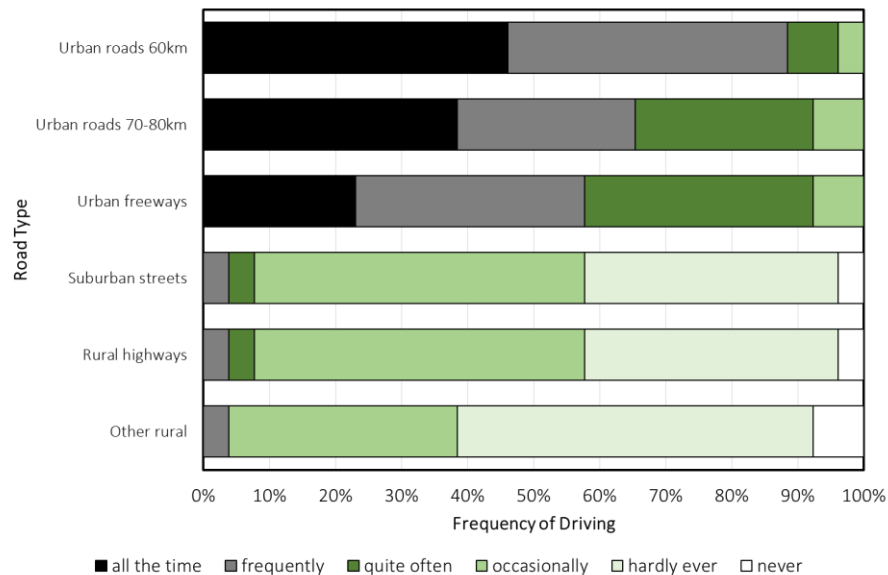


Figure 3-2. Participants' self-reported frequency of driving on different road types.

### 3.3.2. Effects of Image Repetition

Each observer completed were 40 trials involving image repetitions (20 change-present, 20 change-absent). This represented 20 unique images, which were each presented twice.

#### 3.3.2.1. Change-absent trials

Accuracy was at ceiling for change absent trials, regardless of image repetition. Specifically, accuracy on change-absent trials was 99.2% for the first image presentation and 100% for the second image presentation. Due to these values being at ceiling, it was not possible to compare them statistically.

RTs for correct change-absent trials were compared for the first vs. second image presentation using linear GEE with a log link function (as RTs were positively skewed). This comparison indicated no significant difference in change-absent RTs between the first image presentation ( $M = 7122$  ms,  $SE = 395$ ) and second image presentation ( $M = 6886$  ms,  $SE = 517$ ), Wald  $\chi^2(1) = 1.78$ ,  $p = .183$ ,  $B = -0.03$ ,  $SE = 0.03$ , odds ratio (OR) = 0.97, 95% CI OR [0.92, 1.02].

#### 3.3.2.2. Change-present trials

Note that there was only one repetition of each of the 10 change targets, so the image repetition analyses considered the main effect of repetition order (first vs. second) averaging across all types of change targets.

Accuracy for change present trials was 67% for the first image presentation and 72% for the second image presentation. Statistical comparison using binary logistic GEE revealed a significant main effect of image repetition,  $\chi^2(1) = 5.65$ ,  $p = .017$ ,  $B = 0.23$ ,  $SE = 0.10$ , OR = 1.26, 95% CI OR [1.04, 1.53]. That is, participants were significantly more likely to detect changes the second time an image was presented.

RTs for correct change-present trials were compared for the first vs. second image presentation using linear GEE with a log link function (as RTs were positively skewed). This comparison indicated no significant effect of repetition order,  $\chi^2(1) = 13.94$ ,  $p < .001$ ,  $B = -0.07$ ,  $SE = 0.02$ ,  $OR = 0.93$ , 95% CI OR [0.89, 0.97]. Specifically, RTs were shorter for the second image presentation ( $M = 4789$  ms,  $SE = 99$ ) compared with the first image presentation ( $M = 5158$  ms,  $SE = 114$ ).

Visual fixations on the change target were analysed to assess whether patterns of eye movements could explain RT differences between the first and second image repetition. Aspects of fixations that were analysed were: probability of fixating the target; probability of looked-but-failed-to-see errors; time to first fixation (milliseconds); and total dwell time on target (milliseconds). Probability variables were analysed using binary logistic GEE, and time variables were analysed using linear GEE with a log link function (as both were positively skewed). All analyses used image repetition order (first vs. second) as the only factor.

Probability of fixating the target (41% vs. 42% fixated),  $\chi^2(1) = 0.05$ ,  $p = .825$ ,  $B = 0.03$ ,  $SE = 0.14$ ,  $OR = 1.03$ , 95% CI OR [0.78, 1.37], and probability of looked-but-failed-to-see errors (9% in both conditions),  $\chi^2(1) = 0.01$ ,  $p = .941$ ,  $B = -0.03$ ,  $SE = 0.38$ ,  $OR = 0.97$ , 95% CI OR [0.46, 2.04], were not significantly different between the first and second image repetitions.

Total dwell time on the target was also not significantly different in the first image presentation ( $M = 496$  ms,  $SE = 23$ ) compared with the second image presentation ( $M = 474$ ,  $SE = 33$ ),  $\chi^2(1) = 0.48$ ,  $p = .487$ ,  $B = -0.05$ ,  $SE = 0.07$ ,  $OR = 0.96$ , 95% CI OR [0.84, 1.09]. However, time to first fixation was significantly earlier for the second image presentation ( $M = 1495$ ,  $SE = 47$ ) compared with the first image presentation ( $M = 1789$ ,  $SE = 91$ ),  $\chi^2(1) = 16.80$ ,  $p < .001$ ,  $B = -0.18$ ,  $SE = 0.04$ ,  $OR = 0.84$ , 95% CI OR [0.77, 0.91].

Overall the results of the image repetition analyses suggest that participants were more accurate and faster at detecting changes in the second vs. first image repetition. The time to first fixation analyses suggest that priming occurred, as participants were able to shift their gaze to the change target location sooner of the second trial, which seems to account for the differences in RT. Based on this, the subsequent analyses excluded the second presentation of repeated images, so that each participant contributed 200 trials (100 change-present, 100 change-absent) to the main analyses.

### **3.3.3. Change Detection Accuracy**

Accuracy on change-absent trials was at ceiling (99.4% in rural scenes, 99.2% in urban scenes) and so was not included in any statistical analyses.

Among change-present trials, accuracy varied with change target. As shown in Figure 3-3, detection of tree changes was at floor (8% correct), which meant that overall comparisons of performance in urban vs. rural scenes was confounded by target type. As such, urban-rural comparisons were conducted using only targets that appeared in both environments (i.e., road signs, cars, motorcycles), with additional separate analyses for each environment that included safety relevance as a covariate. All analyses used binary logistic GEE.



Within urban scenes, the effect of safety relevance on accuracy was statistically significant,  $\chi^2(1) = 83.62$ ,  $p < .001$ , with participants more likely to detect changes that had higher safety relevance ratings,  $B = 0.65$ ,  $SE = 0.07$ , odds ratio (OR) = 1.92, 95% CI OR [1.67, 2.20]. The main effect of target type was also significant,  $\chi^2(4) = 143.39$ ,  $p < .001$ . Compared to changes involving signs, participants were significantly more likely to detect all other types of changes (see Table 3-1), with the largest effect size for motorcycles.

Within rural scenes, there was a significant main effect of target type on accuracy,  $\chi^2(4) = 163.16$ ,  $p < .001$ . As shown in Table 3-1, compared with changes involving signs participants were less likely to detect changes involving trees, but were more likely to detect changes involving cars, motorcycles and animals. Safety relevance also predicted change detection accuracy in rural scenes, but the effect size was smaller than for urban scenes and only just met the criterion of statistical significance,  $\chi^2(1) = 3.97$ ,  $p = .046$ ,  $B = 0.08$ ,  $SE = 0.04$ , OR = 1.08, 95% CI OR [1.001, 1.17].

Finally, accuracy in urban vs. rural scenes was compared for the three target types that appeared in both environments (road signs, cars, motorcycles). There was a significant main effect of environment,  $\chi^2(1) = 19.22$ ,  $p < .001$ . Compared to rural scenes (92% correct), participants were less likely to detect changes in urban scenes (79% correct),  $B = -0.64$ ,  $SE = 0.13$ , OR = 0.53, 95% CI OR [0.41, 0.68]. There was also a significant main effect of target type,  $\chi^2(2) = 133.92$ ,  $p < .001$ , consistent with the separate urban and rural analyses, but this did not significantly interact with environment,  $\chi^2(1) = 3.77$ ,  $p = .152$ .

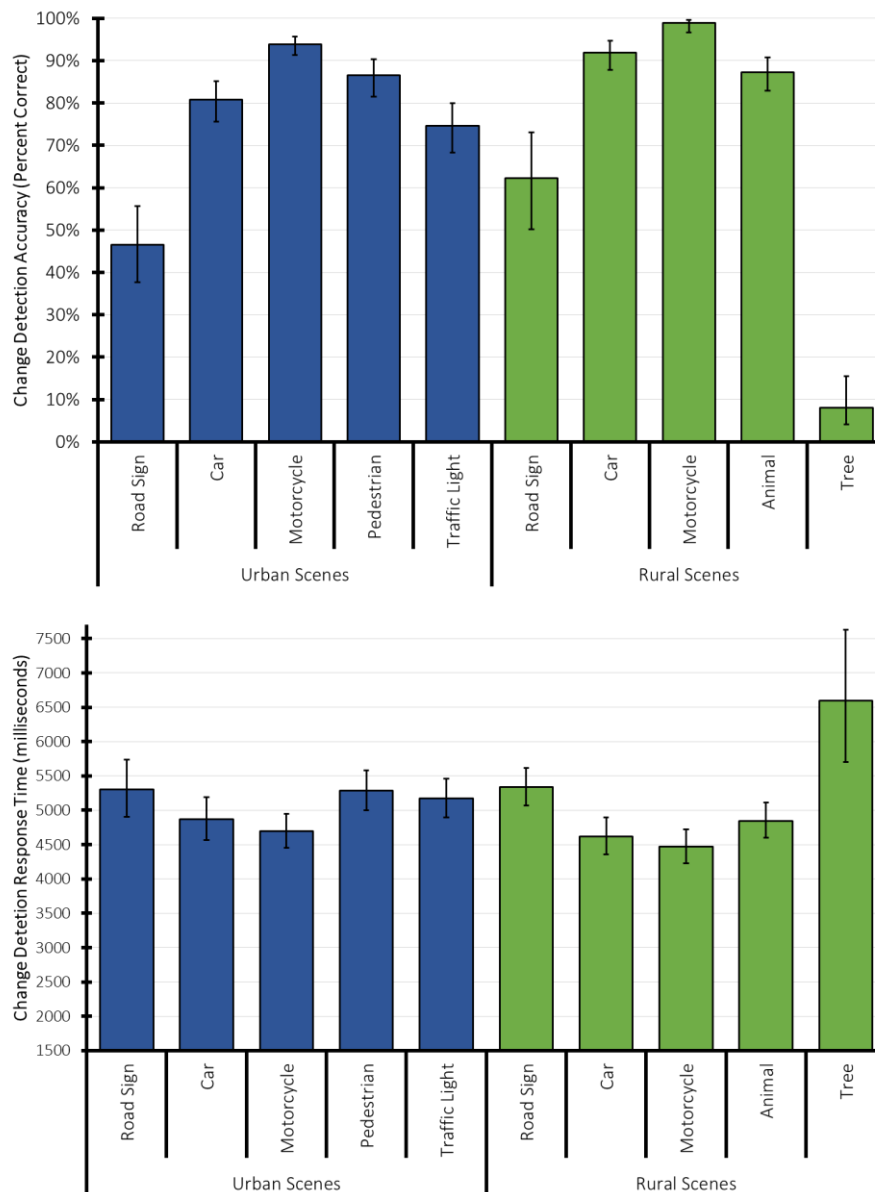


Figure 3-3. Change detection accuracy (top panel) and response time (bottom panel) by driving environment and target type. Error bars represent upper and lower 95% confidence intervals for estimated marginal means within each condition.

Table 3-1

*Statistical comparison of accuracy by target type, within each driving environment*

Target Type	B	SE	Wald $\chi^2$	<i>p</i>	OR	95% CI OR
<i>Urban Scenes</i>						
Traffic Light	0.63	0.20	10.29	.001**	1.88	[1.28, 2.77]
Pedestrian	0.94	0.18	27.00	< .001***	2.56	[1.80, 3.66]
Motorcycle	2.67	0.24	122.86	< .001***	14.49	[9.03, 23.24]
Car	1.71	0.20	71.34	< .001***	5.55	[3.73, 8.26]
Road Sign	-					
<i>Rural Scenes</i>						
Tree	-2.70	0.40	45.81	< .001***	0.07	[0.03, 0.15]
Animal	1.24	0.32	14.69	< .001***	3.44	[1.83, 6.47]
Motorcycle	3.92	0.58	45.38	< .001***	50.41	[16.11, 157.70]
Car	1.96	0.25	63.26	< .001***	7.11	[4.38, 11.52]
Road Sign	-					

*Note.* Road signs were used as the reference category for both urban and rural scene analyses. OR = Odds Ratio. 95% CI = 95% Confidence Interval. \*\**p* < .01, \*\*\**p* < .001.

### 3.3.4. Change Detection Response Time (RT)

RT was analysed for correct trials only, to examine how long participants required to either identify a change (for change-present trials) or determine that the scene was unchanged (for change-absent trials). Trials with RTs over 10 s for change-present trials, or 15 s for change-absent trials, were excluded from the analyses as these represented extreme outliers ( $\leq 1\%$  of responses). All analyses used GEE specifying a normal distribution and a log link function, as RTs were positively skewed. Four analyses were conducted, examining RTs in: urban vs. rural change-absent trials; urban change-present trials by target type; rural change-present trials by target type; and urban vs. rural change-present trials including only the three targets that appeared in both environments (road signs, cars, motorcycles).

Within *change-absent trials*, RTs were compared between urban and rural scenes. The model showed a significant effect of road environment,  $\chi^2(1) = 51.57$ ,  $p < .001$ . The average time required to inspect urban scenes ( $M = 7046$  ms,  $SE = 332$ ) was significantly longer than to inspect rural scenes ( $M = 6623$ ,  $SE = 318$ ),  $B = 0.01$ ,  $SE = 0.01$ ,  $OR = 1.06$ , 95% CI OR [1.05, 1.08].

Within *urban change-present trials*, RTs were analysed with safety relevance as a covariate and target type as a predictor. There was a significant effect of safety relevance,  $\chi^2(1) = 135.09$ ,  $p < .001$ ,  $B = -0.04$ ,  $SE = 0.00$ ,  $OR = 0.96$ , 95% CI OR [0.96, 0.97], with participants responding faster to changes rated as having greater safety relevance. There was a also significant effect of target type,  $\chi^2(4) = 164.01$ ,

$p < .001$  (see Table 3-2). There was a discrepancy between motorised road users and other target types. Specifically, compared to changes involving signs, participants were significantly faster at detecting changes involving cars or motorcycles, but were not significantly faster at changes involving pedestrians or traffic lights, as shown in Figure 3-3.

Within *rural change-present trials*, RTs were also analysed with safety relevance as a covariate and target type as a predictor. The effect of safety relevance was not statistically significant,  $\chi^2(1) = 2.68$ ,  $p = .102$ , but there was a significant effect of target type,  $\chi^2(4) = 82.01$ ,  $p < .001$  (see Table 3-2). As shown in Figure 3-3, the RT results mirrored the pattern obtained for accuracy. Compared to changes involving signs, participants were significantly slower at detecting changes involving trees and significantly faster at detecting changes involving cars, motorcycles or animals.

Finally, RTs were compared between urban vs. rural scenes for the three target types that appeared in both environments (road signs, cars, motorcycles). There was a significant main effect of environment,  $\chi^2(1) = 37.38$ ,  $p < .001$ , with RTs being significantly longer for urban scenes ( $M = 5105$  ms,  $SE = 77$ ) than for rural scenes ( $M = 4803$ ,  $SE = 86$ ),  $B = 0.04$ ,  $SE = 0.02$ ,  $OR = 1.05$ , 95% CI OR [1.004, 1.09]. There was also a significant main effect of target type,  $\chi^2(2) = 53.20$ ,  $p < .001$ , but this did not significantly interact with environment,  $\chi^2(1) = 0.90$ ,  $p = .636$ , consistent with the accuracy results.

Table 3-2

*Statistical comparison of change detection RT by target type, within each driving environment*

Target Type	B	SE	Wald $\chi^2$	$p$	OR	95% CI OR
<i>Urban Scenes</i>						
Traffic Light	-0.03	0.02	1.28	.258	0.98	[0.93, 1.02]
Pedestrian	0.00	0.03	0.02	.886	1.00	[0.94, 1.05]
Motorcycle	-0.12	0.03	20.43	< .001***	0.89	[0.84, 0.93]
Car	-0.09	0.03	9.87	< .001***	0.92	[0.87, 0.97]
Road Sign	-					
<i>Rural Scenes</i>						
Tree	0.21	0.07	10.43	< .001***	1.24	[1.09, 1.41]
Animal	-0.10	0.02	17.50	< .001***	0.91	[0.87, 0.95]
Motorcycle	-0.18	0.03	41.61	< .001***	0.84	[0.79, 0.88]
Car	-0.15	0.03	31.30	< .001***	0.87	[0.82, 0.91]
Road Sign	-					

*Note.* Road signs were used as the reference category for both urban and rural scene analyses. OR = Odds Ratio. 95% CI = 95% Confidence Interval. \*\*\* $p < .001$ .

### 3.3.5. Self-Report Measures

CFQ total scores were computed by summing responses to all items, yielding possible scores of 0 to 100. In the current sample Cronbach's alpha ( $\alpha$ ) was .83 and the range of observed scores was 21-57 ( $M = 39.8$ ,  $SD = 10.2$ ). CFQ scores showed a non-significant small negative correlation with overall change detection accuracy ( $r = -.21$ ,  $p = .307$ ) and a moderate positive correlation with RT ( $r = .39$ ,  $p = .051$ ). Although these trends did not reach statistical significance, they suggest that higher CFQ scores have a small association with poorer change detection performance (i.e., lower accuracy and longer time required to identify changes).

Scores for the DBQ Lapses and Error subscales were computed by summing responses to the items on each scale. This comprised 8 items for the Errors scale (possible scores 0-40) and 7 items for the Lapses scale (possible scores 0-35); one item pertaining to manual transmission cars was excluded because several participants indicated that they exclusively drove automatic transmission cars. For the Errors subscale observed scores were 0-10 ( $M = 4.7$ ,  $SD = 2.5$ ,  $\alpha = .47$ ). For the Lapses subscale observed scores were 2-14 ( $M = 6.9$ ,  $SD = 3.1$ ,  $\alpha = .53$ ). Neither DBQ subscale was significantly correlated with either change detection accuracy (Errors:  $r = -.07$ ,  $p = .749$ ; Lapses:  $r = -.18$ ,  $p = .372$ ) or RT (Errors:  $r = .25$ ,  $p = .216$ ; Lapses:  $r = .16$ ,  $p = .424$ ).

### 3.3.6. Eye Movements: Fixations on Change Targets

Three variables pertaining to fixations on change targets were selected for analysis: probability of fixating the target; probability of looked-but-failed-to-see errors (i.e., failing to detect the change, despite fixating the target); and dwell time on target.

#### 3.3.6.1. Probability of Fixating Target

Probability of target fixation was analysed for all trials, regardless of whether the target was detected, as this represents implicit capture of attention. Binary logistic GEE was used to assess whether probability of fixation differed according to target type and safety relevance, within both urban and rural scenes, with separate analyses for each driving environment.

Within *urban scenes*, there was a significant effect of safety relevance,  $\chi^2(1) = 9.74$ ,  $p = .002$ ,  $B = 0.13$ ,  $SE = 0.04$ ,  $OR = 1.14$ , 95% CI OR [1.05, 1.23], whereby participants were more likely to fixate on targets with higher safety relevance. There was also a significant effect of target type,  $\chi^2(4) = 64.23$ ,  $p < .001$ . Compared to road signs (43% fixated), observers were significantly more likely to fixate both cars (68% fixated;  $\chi^2 = 19.84$ ,  $p < .001$ ,  $B = 1.02$ ,  $SE = 0.23$ ,  $OR = 2.76$ , 95% CI OR [1.77, 4.31]) and motorcycles (65% fixated;  $\chi^2 = 18.12$ ,  $p < .001$ ,  $B = 0.90$ ,  $SE = 0.21$ ,  $OR = 2.46$ , 95% CI OR [1.63, 3.73]), but not pedestrians (40% fixated;  $\chi^2 = 0.26$ ,  $p = .611$ ) or traffic lights (42% fixated;  $\chi^2 = 0.04$ ,  $p = .850$ ).

Within *rural scenes*, there was a significant effect of safety relevance,  $\chi^2(1) = 39.85$ ,  $p < .001$ ,  $B = 0.31$ ,  $SE = 0.05$ ,  $OR = 1.37$ , 95% CI OR [1.24, 1.51]. Similar to urban scenes, in rural scenes participants were more likely to fixate on targets with higher safety relevance, but the effect was even larger for rural scenes. There was also a significant effect of target type,  $\chi^2(4) = 56.48$ ,  $p < .001$ . Compared to road signs (49% fixated), observers were significantly more likely to fixate both cars (64% fixated;  $\chi^2 = 10.18$ ,  $p = .001$ ,  $B = 0.65$ ,  $SE = 0.20$ ,  $OR = 1.92$ , 95% CI OR [1.29, 2.87]) and were less likely to fixate trees (32%

fixated;  $\chi^2 = 7.49$ ,  $p = .006$ ,  $B = -0.70$ ,  $SE = 0.25$ ,  $OR = 0.50$ , 95% CI OR [0.30, 0.82]). Probability of fixating motorcycles (51% fixated;  $\chi^2 = 0.25$ ,  $p = .618$ ) and animals (39% fixated;  $\chi^2 = 2.94$ ,  $p = .086$ ) was not significantly different to signs.

Finally, an additional analysis comparing probability of fixating the target between urban and rural scenes (for sign, car and motorcycle targets only) revealed no significant effect of driving environment on probability of target fixation,  $\chi^2(1) = 1.42$ ,  $p = .233$ . The effect of target type was also significant, consistent with the analyses conducted separately for urban and rural scenes.

### 3.3.6.2. Probability of Looked-But-Failed-To-See Errors

This analysis focused on the probability of *failing* to detect a change despite having fixated on the target. As with other analyses, comparisons examining the effects of target type and safety relevance were made separately for urban and rural scenes, followed by a direct urban vs. rural comparison.

Within *urban scenes*, participants experienced looked-but-failed-to-see errors on 8% of all trials in which they fixated the target. There were significant effects of both safety relevance,  $\chi^2(1) = 12.11$ ,  $p = .001$ ,  $B = -0.48$ ,  $SE = 0.14$ ,  $OR = 0.62$ , 95% CI OR [0.47, 0.81], and target type,  $\chi^2(4) = 52.52$ ,  $p < .001$ . Observers were less likely to make looked-but-failed-to-see errors for targets with higher safety relevance ratings, regardless of target type. As shown in Table 3-3, looked-but-failed-to-see errors were most common when the target was a road sign, and were significantly less likely in all other conditions.

Table 3-3

*Probability of looked-but-failed-to-see errors in urban scenes, by target type*

Target Type	M	B	SE	Wald $\chi^2$	p	OR	95% CI OR
Traffic Light	8%	-0.97	0.44	4.97	.026*	0.38	[0.16, 0.89]
Pedestrian	1%	-2.98	1.02	8.60	.003*	0.05	[0.01, 0.37]
Motorcycle	<1%	-3.91	0.93	17.68	< .001***	0.02	[0.003, 0.12]
Car	5%	-1.43	0.36	15.47	< .001***	0.24	[0.12, 0.49]
Road Sign	18%	-					

*Note.* Road signs were used as the reference category. OR = Odds Ratio. 95% CI = 95% Confidence Interval. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Within *rural scenes*, 10% of trials involved looked-but-failed-to-see errors; however, this was inflated by the fact that participants experienced looked-but-failed-to-see errors on 71% of trials in the tree condition, compared to 0% for motorcycles, 2% for animals, 5% for vehicles and 17% for signs. Inspection of the data revealed that target type was confounded with both safety relevance ratings and probability of looked-but-failed-to-see errors, which precluded the possibility of reliable statistical analysis. Binary logistic GEE with safety relevance as the only covariate (i.e., target type was omitted from the model) revealed no significant effects,  $\chi^2(1) = 2.27$ ,  $p = .132$ , suggesting that in rural scenes target type was a better predictor of looked-but-failed-to-see errors than safety relevance of that target.

Finally, an additional analysis comparing probability of looked-but-failed-to-see errors between urban and rural scenes (for sign, car and motorcycle targets only) revealed a significant main effect of driving environment,  $\chi^2(1) = 7.49$ ,  $p = .006$ , whereby looked-but-failed-to-see errors were slightly but significantly more common in urban (5%) vs. rural (3%) scenes,  $B = 0.62$ ,  $SE = 0.23$ ,  $OR = 1.86$ , 95% CI  $OR [1.19, 2.89]$ . The effect of target type was also significant, consistent with the analyses conducted separately for urban and rural scenes.

### 3.3.6.3. Dwell Time on Target

Dwell time indicates the relative difficulty of identifying targets that are fixated; longer dwell times indicate the participant requires more time to cognitively process the target. The analyses included only correct trials in which the participant fixated the target. As with other measures, separate analyses were conducted for urban and rural scenes, followed by a direct urban vs. rural comparison.

Within *urban scenes*, there were significant effects of both safety relevance,  $\chi^2(1) = 9.47$ ,  $p = .002$ ,  $B = -0.06$ ,  $SE = 0.18$ ,  $OR = 0.95$ , 95% CI  $OR [0.91, 0.98]$ , and target type,  $\chi^2(4) = 54.76$ ,  $p < .001$ . Dwell times were shorter on targets with higher safety relevance. As shown in Table 3-4, the results for dwell time mirrored the patterns for change detection accuracy: compared with road signs dwell times were significantly shorter for all other target types, with the effect being largest for motorcycles.

Within *rural scenes*, there was a significant effect of safety relevance,  $\chi^2(1) = 22.14$ ,  $p < .001$ ,  $B = 0.09$ ,  $SE = 0.02$ ,  $OR = 1.09$ , 95% CI  $OR [1.05, 1.13]$ , but the effect was in the opposite direction to that found in rural scenes: targets with higher safety relevance were associated with longer dwell times. This is likely a statistical artefact, due to the confound between target type and safety relevance, as the zero-order correlation between safety relevance and dwell time trended in the opposite direction. There was also a significant effect of target type,  $\chi^2(4) = 180.33$ ,  $p < .001$ , with considerable variations in dwell time between targets, as shown in Table 3-4. Compared to road signs, observers spent significantly less time looking at animals, motorcycles and cars, but more time looking at trees.

Finally, dwell times were compared between urban and rural scenes, for trials where the target was a road sign, car or motorcycle. This analyses revealed significant effects of target type, consistent with the separate urban and rural analyses, but no effect of driving environment,  $\chi^2(1) = 0.07$ ,  $p = .797$ .

Table 3-4

*Average dwell time (in milliseconds) on the change target, by target type and driving environment*

Target Type	<i>M</i>	<i>B</i>	<i>SE</i>	Wald $\chi^2$	<i>p</i>	OR	95% CI OR
<i>Urban Scenes</i>							
Traffic Light	655	-0.20	0.08	5.71	.017*	0.82	[0.70, 0.97]
Pedestrian	510	-0.45	0.08	33.44	< .001***	0.64	[0.55, 0.74]
Motorcycle	418	-0.65	0.09	47.37	< .001***	0.52	[0.45, 0.63]
Car	577	-0.32	0.07	23.04	< .001***	0.73	[0.64, 0.83]
Road Sign	786	-					
<i>Rural Scenes</i>							
Tree	1606	0.54	0.22	5.89	.015*	1.72	[1.11, 2.67]
Animal	328	-1.05	0.10	108.71	< .001***	0.35	[0.29, 0.43]
Motorcycle	428	-0.78	0.07	113.51	< .001***	0.46	[0.40, 0.53]
Car	667	-0.34	0.08	16.95	< .001***	0.72	[0.61, 0.84]
Road Sign	933						

*Note.* Road signs were used as the reference category. OR = Odds Ratio. 95% CI = 95% Confidence Interval. \* $p < .05$ , \*\*\* $p < .001$ .

### 3.3.7. Eye Movements: Non-Target Fixation Patterns

To examine scanning patterns more generally, several aspects of eye movements were compared between urban and rural change-absent trials. These measures included the average number and duration of fixations made each trial, as well as the probability of fixating specific regions of interest within the scene and dwell times on those regions. Five interest area (IA) regions were defined on each image: the road itself; off-road left; off-road right; horizon (where road meets sky); and sky.

As shown in Table 3-5, observers made more significantly more fixations per trial, but significantly shorter fixations, when viewing urban scenes compared to rural scenes. There were also differences in where observers fixated: the probability of fixating all five IAs was significantly higher in urban vs. rural scenes. Dwell times (measured as a proportion of the total dwell time for the trial) were significantly longer on the road IA for rural vs. urban scenes, but were significantly longer on the off-road-right and sky IAs for urban vs. rural scenes. This indicates that when viewing rural scenes, participants mostly focused their attention on the road itself, whereas in urban scenes they devoted more time to searching other areas of the scene.



Table 3-5

*Patterns of eye movements in change-absent trials, comparing between driving environments*

Measure	Urban	Rural	Difference		Comparison
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M</i>	95% CI	
Average fixations per trial	15.4 (5.5)	13.6 (4.8)	1.8	[1.3, 2.2]	$t(25) = 7.62$ , $p < .001^{***}$ , $d = 1.49$
Average fixation duration	315 (52)	332 (52)	17	[12, 23]	$t(25) = 6.26$ , $p < .001^{***}$ , $d = 1.23$
Probability of fixation:					
IA: Road	94% (10%)	92% (11%)	2%	[0%, 3%]	$t(25) = 2.34$ , $p = .028^*$ , $d = 0.46$
IA: Off-road left	92% (11%)	82% (14%)	10%	[7%, 13%]	$t(25) = 7.08$ , $p < .001^{***}$ , $d = 1.39$
IA: Off-road right	89% (6%)	75% (8%)	14%	[11%, 17%]	$t(25) = 10.56$ , $p < .001^{***}$ , $d = 2.07$
IA: Horizon	92% (6%)	86% (12%)	6%	[3%, 10%]	$t(25) = 3.66$ , $p = .001^{**}$ , $d = 0.72$
IA: Sky	84% (8%)	52% (15%)	33%	[29%, 37%]	$t(25) = 17.06$ , $p < .001^{***}$ , $d = 3.35$
Dwell time (% of trial)					
IA: Road	29% (9%)	34% (13%)	5%	[2%, 07%]	$t(25) = 3.64$ , $p = .001^{**}$ , $d = 0.71$
IA: Off-road left	29% (6%)	28% (6%)	1%	[0%, 03%]	$t(25) = 1.61$ , $p = .120$ , $d = 0.32$
IA: Off-road right	26% (4%)	23% (4%)	3%	[1%, 05%]	$t(25) = 3.43$ , $p = .002^{**}$ , $d = 0.67$
IA: Horizon	32% (6%)	31% (7%)	1%	[-1%, 04%]	$t(25) = 1.03$ , $p = .312$ , $d = 0.20$
IA: Sky	16% (5%)	10% (4%)	6%	[5%, 08%]	$t(25) = 10.96$ , $p < .001^{***}$ , $d = 2.15$

Note. 95% CI = 95% Confidence Interval. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

### 3.4. Discussion

The aim of the current study was to examine drivers' change detection ability in urban and rural driving scenes, for a range of objects that vary in their contextual safety relevance. All participants were experienced, fully-licensed drivers who drove at least weekly and were familiar with the locations depicted in the stimulus images, although they reported driving considerably more frequently in urban areas compared to rural roads. The results revealed several aspects of change detection performance, including accuracy, RT and eye movements, vary as a function of driving environment, target type, and the safety relevance of the change.

#### 3.4.1. *Effects of Driving Environment: Urban vs. Rural*

When directly comparing performance in urban and rural scenes, with target type and context matched between environments, participants were significantly more accurate and faster at detecting changes in rural scenes compared with urban scenes. Participants were also more likely to exhibit “looked-but-failed-to-see” errors, whereby they fixated the target but failed to consciously detect and report the change, although the effect size was relatively small (3% vs. 5%). These differences are most likely attributable to the fact that urban scenes involve greater visual clutter and complexity. To our knowledge, no previous research has compared change detection in urban and rural scenes in the same way as the current study. However, these findings are consistent with research on visual crowding (Whitney & Levi, 2011). Also, it is worth noting that participants were significantly more familiar with urban driving, and drove regularly in the areas depicted in the urban scenes, whereas they reported significantly less exposure to rural driving. In this regard, the results are consistent with previous findings regarding the effects of familiarity on change blindness (e.g., Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007), which indicate that drivers exhibit greater change blindness in familiar situations.

Despite the slight increase in looked-but-failed-to-see errors in urban scenes, there was no difference in the probability of fixating targets, or total dwell time on targets, when comparing urban and rural scenes. Analyses of eye movements in change-absent trials suggest this could be because participants adopted different scanning patterns when viewing urban scenes, to maximise their likelihood of detecting target objects in cluttered urban environments. Specifically, when viewing urban scenes participants made more fixations of shorter average duration, and distributed their fixations more broadly throughout the scene, whereas when viewing rural scenes participants made fewer longer fixations and focused predominantly on the road itself. This is consistent with research on eye movements in driving, which has found that experienced drivers adapt their scanning patterns based on situational demands (Chapman & Underwood, 1998; Falkmer & Gregersen, 2005; Underwood, 2007; Underwood et al., 2002).

#### 3.4.2. *Effects of Change Safety Relevance*

In addition to the differences that emerged from the direct comparison of urban and rural scenes, the analyses regarding safety relevance of changes revealed different patterns for the two driving environments. Specifically, the effects of the safety relevance of the change were larger and more consistent for the urban scenes. Changes with higher safety relevance were associated with higher

accuracy, shorter RT, increased probability of fixating the target, reduced probability of looked-but-failed-to-see errors, and shorter dwell times. Taken together, these findings suggest that changes with greater safety relevance are more effective at capturing drivers' implicit attention (i.e., probability of fixation) and then, due to their relevance, are processed into conscious awareness. These findings are consistent with previous change detection research, which has consistently revealed that observers are more efficient at changes that are more central to interpreting the scene (Rensink et al., 1997) and those that have greater personal or task relevance (Galpin et al., 2009; Lee et al., 2007; Marchetti et al., 2006; Mueller & Trick, 2013; Shinoda et al., 2001; Velichkovsky et al., 2002; Zhao et al., 2014).

In contrast to the results observed in urban scenes, the effects of safety relevance on detection of changes in rural scenes was considerably less consistent. Safety relevance of the change had only a marginally significant effect on change detection accuracy in rural scenes and did not predict RT or looked-but-failed-to-see errors. The only measure that was clearly affected in the expected direction was probability of fixating the target, in that drivers were more likely to fixate targets with higher safety relevance. One possibility is that these inconsistent effects are linked to the task demands and resulting performance differences between urban and rural scenes. That is, urban scenes were more cognitively demanding to process and so observers preferentially focused on aspects of the scene that appeared to have greater relevance. Rural scenes were easier to process, which meant that participants had the capacity to process change targets that had lower safety relevance.

### ***3.4.3. Effects of Target Type***

Beyond the effects of change safety relevance, there were also significant effects of target type on change detection performance, especially for trees and signs.

Change detection performance was at floor for changes involving trees, with most participants failing to detect all of the tree-related changes. Participants were also less likely to fixate on trees compared to other target types and were substantially more likely to exhibit looked-but-failed-to-see errors on the occasions when they did fixate trees. These patterns cannot be wholly explained by safety relevance, as target position was systematically manipulated so that half of the trees appeared directly next to the road (where they posed a potential hazard in the event of an emergency). However, the fact that drivers overlooked changes to roadside foliage is consistent with previous research on risk perception, which found that participants consistently overlook subtle roadside features that increase the hazardousness of driving on a particular road (Charlton et al., 2014).

When changes involved signs, participants were significantly less efficient at change detection compared to all other types (excluding trees). In both urban and rural scenes, participants were less accurate and exhibited longer dwell times for sign changes, compared to other types of changes. These results are consistent with previous research, which found that participants commonly exhibit change blindness when changes involve road signs (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007). One commonality across the non-sign, non-tree target types is that they are all objects that could plausibly change: cars, motorcycles, pedestrians and animals are all mobile, whereas traffic lights have a fixed position but update dynamically. As such, one possibility may be that participants were

preferentially attending to aspects of the scene that are most likely to change in a real driving environment. This is consistent with the fact that changes to trees were almost never detected.

Another explanation is that participants preferentially attend to elements within the scene that are potentially dangerous. This is supported by RT, probability of fixation, and looked-but-failed-to-see error analyses. Specifically, changes involving pedestrians and traffic lights were not significantly different from sign changes in terms of RT, probability of target fixation, and looked-but-failed-to-see errors. In contrast, when changes involved cars, motorcycles, or animals, participants exhibited shorter RTs, increased probability of fixating the target, and reduced probability of looked-but-failed-to-see errors. The key difference between cars, motorcycles and animals on the one hand, and pedestrians and traffic lights on the other hand, is that the former category are more likely to cause damage in the event of a collision. (Keeping in mind that several of the animal targets were kangaroos, which pose a particular threat to drivers in the Canberra region.)

#### **3.4.4. Summary**

Overall the results of Experiment 1 indicate that change detection efficiency is affected by several variables, including the driving environment in which the change occurs, the safety relevance on the change, and the type of object changed. Specifically, drivers are more efficient at detecting changes to other road users or potential hazards, such as animals near the roadside, as well as changes with higher safety relevance. Drivers are also better at detecting changes in rural scenes compared to urban scenes, which is likely because there is less visual clutter in rural areas, but could also reflect the fact that urban areas are more familiar (which has been demonstrated to exacerbate change blindness).

Most notably, all of the change targets in the current study were potentially driving relevant, in that they were road users or roadside objects. The results therefore demonstrate that not all “driving relevant” changes are equal, which has implications for previous research that used broad categories to define relevant vs. irrelevant images.

A final point worth noting is the fact that the self-report measures of cognitive failures and driving-related errors and lapses were not significantly associated with change detection ability. This is suggestive of “*change blindness blindness*”, which refers to the fact that observers do not have a good understanding of their own change detection ability and commonly under-estimate their susceptibility to change blindness (Beck et al., 2007). In the context of driving, this could be problematic if drivers are not aware of precisely how difficult it is to detect changes, especially for changes involving road signs. Two main avenues are available for addressing this issue. First, driver education programs could aim to raise awareness of change blindness, specifically highlighting the types of changes that drivers are most likely to have trouble detecting. (Note that some driver education programs do mention change blindness and/or inattention blindness, but often use generic examples rather than focusing on specifics of when these phenomena are likely to occur on the road.) Second, road sign design and placement should be rigorously evaluated and changed where appropriate, so that redundant signs can be eliminated and safety-critical signs can be redesigned to better capture drivers’ attention.

## 4. Experiment 2

### 4.1. Background and Rationale

Driver sleepiness is a causal factor in approximately 15-30% of all crashes (Åkerstedt, 2000; Connor, 2009; Horne & Reyner, 1995). Crash outcome is often severe, with drivers who are sleepy being at an almost six fold increase in the odds of having an injury-involved crash (Herman et al., 2014). Sleep-related crashes are most commonly characterised by the vehicle drifting out of the driving lane and colliding with an object in plain sight, and there are often no signs of braking or attempted avoidance manoeuvres by the sleepy driver (Horne & Reyner, 1995). These types of crashes most often occur on high speed roads in rural environments. Extreme sleepiness (having fallen asleep or had to stop driving), which may result in out-of-lane events while driving, is experienced by approximately 8-9% of drivers every month (Philip et al., 2010; Sagberg, 1999). However, the majority of drivers experience some degree of sleepiness on some occasions (Armstrong et al., 2013). To date there has been little attempt to understand the driving impairment experienced due to slight sleepiness prior to the point of experiencing a micro sleep and/or having an out-of-lane incident.

Although driver sleepiness makes up a greater proportion of total crashes in high speed ( $\geq 100$  km/h) zones, a recent analysis of Queensland crash data reported that over 40% of sleep-related crashes occur in low speed zones ( $\leq 60$  km/h; Filtiness et al., under review). Similarly, in a self-report survey of ACT and NSW drivers who had had a sleep-related driving incident, 25% reported that this incident occurred in a residential area with speed limit of 50 km/h or less and a further 30% reported an incident occurring on roads with speed limits between 50 and 80 km/h (Armstrong et al., 2013). To date, the majority of driver sleepiness research has focused on understanding driver sleepiness during rural or motorway driving (e.g., Filtiness et al., 2012; Hallvig et al., 2013; Philip et al., 2005), with little attempt to specifically investigate low speed sleep-related crashes.

Sleepiness can cause a range of deficiencies which have potential to subtly impair driving performance. For example, sleepiness slows reaction time, impairs decision making ability, and reduces vigilance (Jackson et al., 2013), all of which are essential skills for safe driving. Furthermore, sleepiness impairs complex cognitive processes, such as information processing and planning skills (Horne, 2012), as well as reducing the ability to complete dual task paradigms (Haavisto et al., 2010). Division of attention and forward planning are both skills vital for the detection of and response to hazards while driving. Another skill necessary for accurate hazard detection and response is visual scanning. Recently it has been noted that sleep deprivation leads to decreased oculomotor function (De Gennaro et al., 2000; Fransson et al., 2008), which impairs visual search performance (De Gennaro et al., 2001). These subtle impairments may interact with each other to impact driving performance prior to the moment of falling asleep and exiting the road. Further, it may be argued that these skills (i.e., rapid decision making and reactions) are relatively more important in urban driving compared with rural driving.

Although sleep loss impairs several skills that are vital to safe driving in urban environments, and drivers commonly report driving while sleepy in urban areas (e.g., Armstrong et al., 2013), no previous research has shown a direct relationship between sleepiness and driving safety in urban areas. The current work considers the impact of sleep loss on the ability to detect changes in driving scenes. Sleep was restricted

to a level expected to invoke sleepiness but not so extreme as to expect participants would fall asleep during the study. The study was designed to assess whether sleep loss impacts change detection for driving scenes, and whether the impact of sleepiness on change detection varies as a function of the driving environment or the type of change that occurred, as both of these factors were found to substantially influence change detection efficiency in Experiment 1.

## 4.2. Method

### 4.2.1. Participants

Twenty-two fully-licensed drivers (15 female, 8 male) aged 20-29 years ( $M = 22.4$ ,  $SD = 2.4$ ) provided informed consent and participated voluntarily. Participants were offered AUD\$50 compensation for their time, plus an additional allowance to cover travel expenses to attend the sleep restriction session. All participants had normal or corrected-to-normal visual acuity, as measured using a near vision chart, and drove at least once a week within the Canberra region.

Participants were pre-screened to ensure they met relevant inclusion criteria for participating in a sleep restriction study. Specifically, participants were required to be regular 7-8 hour/night sleepers who did not take regular naps, suffer from extreme daytime sleepiness, or have any sleep disorders. Participants were excluded if they smoked, drank alcohol daily, and/or they consumed five or more high-caffeine drinks per day.

Ethical aspects of the research were approved by the Australian National University Human Research Ethics Committee (protocol 2014/458).

### 4.2.2. Apparatus

Visual stimuli were presented on a 27" Apple iMac desktop computer. An Eyelink 1000 eye-tracker, with a reported spatial accuracy within  $0.25$ - $0.5^\circ$ , was used to monitor eye movements at a temporal frequency of 1000 Hz. Head position was fixed using a chinrest with a viewing distance of 95 cm, yielding a display area of  $30.3^\circ \times 19.4^\circ$  visual angle. Stimulus presentation and data acquisition were controlled via SR Research Experiment Builder.

BodyMedia SenseWear Armbands were used to monitor participants' sleep and waking activity during the three days preceding each testing session. SenseWear Armbands are wearable physiological monitoring devices that record several parameters; of particular relevance to the current study it records time spent lying down as well as sleep duration and efficiency.

### 4.2.3. Stimuli

There were two matched sets of experimental stimuli (stimulus sets A & B), one for each change detection session. The images used within sets A & B were different (but matched for difficulty of change detection) to control for image repetition priming effects demonstrated in Experiment 1.

Each stimulus set included 80 image pairs depicting driving scenes: 20 urban change-present pairs, 20 rural change-present pairs, 20 urban change-absent pairs and 20 rural change-absent pairs. In change-



absent pairs the two images displayed were identical, whereas in change-present pairs one of the images was edited to add, remove or alter a single driving-relevant target.

Each image subtended  $23.0^{\circ} \times 17.5^{\circ}$  visual angle and was taken using a digital camera mounted on the dashboard of a station wagon. Urban images were taken in central Canberra (civic, inner north, Parliamentary Triangle) and rural images were taken on rural roads in surrounding regions.

Within both the urban and rural environments, five types of change targets were used. In the urban scenes the change targets were either cars, motorcycles, road signs, traffic lights or pedestrians, with four image pairs for each category. In the rural scenes the change targets were either cars, motorcycles, road signs, trees or animals, again with four image pairs for each category.

To develop the matched stimulus sets, we analysed the Study 1 data for each of the 100 change-present trials, comparing RT and accuracy (averaged across all participants) for each trial within the 10 different stimulus categories (i.e., urban/car, urban/motorcycle, urban/sign, urban/pedestrian, urban/traffic light, rural/car, rural/motorcycle, rural/sign, rural/tree, rural/animal). The purpose of this was to identify change detection trials that had similar levels of difficulty. Where two trials had similar difficulty, the image pairs used in one trial were assigned to stimulus set A and the other image pair was assigned to stimulus set B. Some trials appeared to be outliers, in that change detection performance was unusually good (high accuracy, low RT) or poor (low accuracy, high RT) for that target category, and these image pairs were excluded. This resulted in 40 change-present image pairs for each stimulus set (80 total), with four repetitions of each target type.

For consistency, we also included only 40 change-absent image pairs (20 urban, 20 rural) in each stimulus set, and images in each stimulus set were matched on RT and accuracy.

Image pairs were presented using a “flicker” sequence, in which one image was presented for 500 ms, followed by a 500 ms blank grey screen, followed by the second image for 500ms and then another 500 ms blank. The cycle of alternating images and blanks continued until the participant responded, or for 30 s, whichever occurred first. Participants were instructed to decide as quickly as possible whether a change occurred and then immediately press the space bar to register their decision. They were then prompted to report whether a change occurred and, if applicable, the change target. If participants failed to respond within 30 s the program automatically proceeded to a response screen that asked them to indicate whether a change occurred. Available response options included “yes” and “no” for whether a change occurred, and “vehicle”, “motorcycle”, “bicycle”, “person”, “animal”, “tree”, “building”, “sign”, and “traffic light” for change target.

Change-present trials were considered “correct” if the observer correctly identified the change target, but were considered “incorrect” if they reported no change or failed to select the correct change target. Change-absent trials were considered “correct” if the observer reported no change, and were considered “incorrect” if they indicated a change occurred.

#### **4.2.4. Self-Report Measures**

During the introductory session participants completed a demographics questionnaire. This included questions confirming participants met the screening criteria and description of their usual driving

exposure and behaviour. In addition, the Epworth Sleepiness Scale (ESS; Johns, 1991) was used to identify if any participants experienced excessive day time sleepiness ( $ESS > 12$ ).

For the three nights prior to each experimental study session participants were required to keep daily sleep diaries of their bed time, estimated sleep onset, night time wakings, and morning awakening and rising times. These self-report measures were considered alongside the objective SenseWear Armband recording of sleep.

Participants were asked report their subjective sleepiness on the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990) at the start and end of each study session. The KSS measures subjective sleepiness at a given point in time on a 9-point scale: (1) extremely alert; (2) very alert; (3) alert; (4) rather alert; (5) neither alert nor sleepy; (6) some signs of sleepiness; (7) sleepy, no effort to stay awake; (8) sleepy, some effort to stay awake; and (9) very sleepy, great effort to keep awake, fighting sleep.

#### **4.2.5. Procedure**

Participants attended the lab for three 30-minute sessions, which were held on separate days at least three days apart. These comprised one introductory session followed by two change detection test sessions. All sessions were scheduled on weekday afternoons, at either 1400h or 1445. Participants completed all three sessions at the same time (i.e., a given participant would complete all three sessions at 1400h, or all three sessions at 1445h). One change detection session was completed following a normal night of sleep (Normal Sleep; NS) and one after sleep restriction (SR) to five hours, which was achieved by instructing participants to delay their bed-time by 3 hours on the night before the SR session. The order of change detection sessions (i.e., NS vs. SR first) was counterbalanced between participants. Further, presentation of stimulus sets was counterbalanced such that half the participants received stimulus set A in the normal sleep session and B in the sleep restriction session, whereas the other half received set B in the normal sleep session and A in the sleep restriction session.

In the introductory session participants provided written informed consent, completed the background questionnaires, and were given the SenseWear Armband and sleep diaries with instructions on how to use the armband and record their sleep.

In the two change detection sessions, participants provided their sleep diary and armband to a research assistant, who checked the data to ensure compliance with the required hours of sleep. Once this was confirmed, the participant completed the KSS to indicate their pre-task subjective sleepiness. Participants then completed the change detection task, which included 10 practice trials before the main change detection task, with a break halfway through the task. The eye-tracker was calibrated for each participant's gaze at the beginning of the study, and recalibrated after the break, with drift checks conducted at the start of each trial to ensure accurate gaze tracking was maintained. Finally, after completion of the change detection task, there was a second administration of the KSS to measure post-task subjective sleepiness.

#### **4.2.6. Data Analysis**

Paired *t*-tests were used to compare sleep duration, subjective sleepiness, and oculomotor behaviour between the Normal Sleep (NS) and Sleep Restriction (SR) conditions.



For change-present trials accuracy, RT, target fixations and dwell time were each analysed using RM-ANOVA with two within-subjects factors: Sleep Condition (2 levels: NS, SR) and Change Target (5 levels; Urban: sign, car, motorcycle, pedestrian, traffic light; Rural: sign, car, motorcycle, animal, tree). Urban and Rural results were first analysed separately because the types of change targets varied between the environment conditions (i.e., pedestrians and traffic lights changed in urban scenes only; animals and trees changed in rural scenes only).

To compare between Urban and Rural environments only the three target types that appeared in both environments (i.e., signs, cars, and motorcycles) were considered. RM-ANOVAs were undertaken with three within-subjects factors: Sleep Condition (2 levels: NS, SR); Driving Environment (2 levels: Urban, Rural) and Change Target (3 levels: sign, car, motorcycle).

For change-absent trials, accuracy and RT were compared using RM-ANOVA with two within-subjects factors: Sleep Condition (2 levels: NS, SR) and Driving Environment (2 levels: Urban, Rural).

All statistical analyses were conducted using SPSS 21.0 statistical software. An alpha level of .05 was used to determine statistical significance. For ANOVAs, post hoc pairwise comparisons were conducted using Bonferroni tests. To supplement the interpretation of the results, partial  $\eta^2$  was used as an estimate of effect size. Where Mauchly's test indicated that the assumption of sphericity had been violated, degrees of freedom were corrected using Huynh–Feldt estimates of sphericity, and epsilon ( $\epsilon$ ) values are listed accordingly.

### 4.3. Results

#### 4.3.1. Participants

Twenty-two participants were recruited for the study. Two participants dropped out after the first study session. Results are presented for 20 participants (14 female).

All participants were aged 20-30 years ( $M = 22.35$ ,  $SD = 2.37$ , range 20-29). Participants were frequent drivers ( $M = 7.98$  hours/week,  $SD = 8.94$ , range 1-42), covering a mean of 263 km per week ( $SD = 258$ , range 30-1000). One participant wore lenses to correct their vision. No participants suffered from excessive daytime sleepiness, defined as ESS scores above 12 ( $M = 4.3$ ,  $SD = 2.5$ , range 0-11). Additional participant characteristics are presented in Table 4-1.

Due to dropouts and scheduling issues, there was some inconsistency in the counterbalancing of image sets used. Following a normal night of sleep (NS session), 7 participants viewed image set A, whereas 13 participants viewed image set B in their NS session.

Table 4-1

*Participant characteristics*

Characteristic	
State/Territory obtained driving licence	
ACT	14
NSW	5
NT	1
Months held a driving licence	53.5 ( $SD = 25.3$ , range 21-120)
Highest Education level	
High School	12
Undergraduate Degree	6
Postgraduate Degree	1
Missing information	1

#### 4.3.2. Sleep Duration

Participants slept for an average of 494 minutes ( $SEM = 8$  min) prior to the NS condition. Sleep was restricted to an average of 303 minutes ( $SEM = 7$  min) during the SR condition. This is a significant reduction,  $t(19) = 8.73$ ,  $p < .001$ .

#### 4.3.3. Subjective Sleepiness

Participants were asked to rate their sleepiness using the KSS when they first arrived and just before they left their study session. Participants felt significantly sleepier following sleep restriction, results are presented in Table 4-2.

Table 4-2

*Mean (and SEM) subjective sleepiness on arrival and before leaving for the two study sessions.*

	Normal Sleep	Sleep Restriction	Comparison
KSS on arrival	3.5 (0.3)	6.4 (0.3)	$t(19) = 10.05$ , $p < .001$
KSS before leaving	3.8 (0.3)	6.6 (0.3)	$t(19) = 15.66$ , $p < .001$

#### 4.3.4. Change Detection Accuracy

##### 4.3.4.1. Rural Scenes

Sleep condition did not influence change detection accuracy in rural scenes,  $F(1,19) = 2.15$ ,  $p = .159$ , partial  $\eta^2 = .10$ . Accuracy for detecting changes involving animate targets (i.e., cars, animals, and motorcycles) was near perfect (>93%) regardless of sleep condition. Accuracy for detecting changes to static objects was lower, approximately 80% for signs and <15% for trees. This is reflected by a significant main effect of change target on accuracy,  $F(2.3,44.5) = 174.28$ ,  $p < .001$ , partial  $\eta^2 = .90$ ,  $\epsilon = .59$ . Pairwise comparisons confirmed that accuracy for changes involving trees was significantly lower than all other target types [ $p < .001$ ]. Accuracy for sign changes was significantly lower than for motorcycles [ $p = .016$ ]. There was no significant interaction between sleep condition and change target,

$F(2.6, 49.1) = 0.25$ ,  $p = .830$ , partial  $\eta^2 = .13$ ,  $\epsilon = .65$ . Change detection accuracy in rural scenes is presented in Figure 4-1.

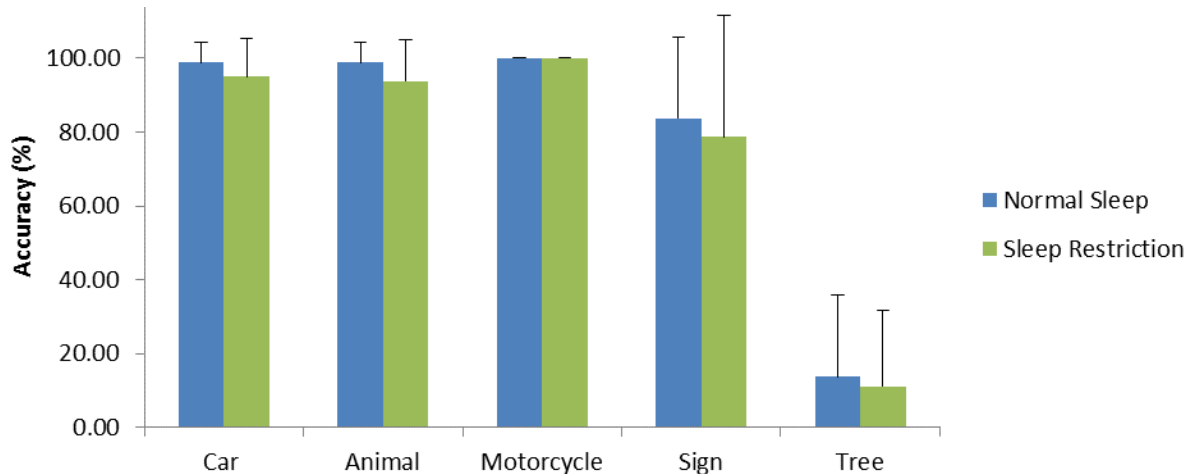


Figure 4-1. Change detection accuracy in rural driving scenes, by change target and sleep condition. Error bars represent standard deviation.

#### 4.3.4.2. Urban Scenes

There was no main effect of sleep condition on change detection accuracy in urban scenes,  $F(1, 19) = 1.51$ ,  $p = .234$ , partial  $\eta^2 = .74$ . However, there was a significant main effect of change target type on accuracy,  $F(3.2, 61.3) = 59.49$ ,  $p < .001$ , partial  $\eta^2 = .76$ ,  $\epsilon = .83$ . Pairwise comparison identified that accuracy changes for signs was significantly lower than for all other target types [ $p < .001$ ]. Participants were significantly more accurate at detecting changes involving motorcycles compared to traffic lights [ $p = .001$ ]. Accuracy for cars, people and motorcycles was near ceiling (>93%). There was no significant interaction between sleep condition and object type,  $F(4, 76) = 0.81$ ,  $p = .521$ , partial  $\eta^2 = .04$ . Results are presented in Figure 4-2.

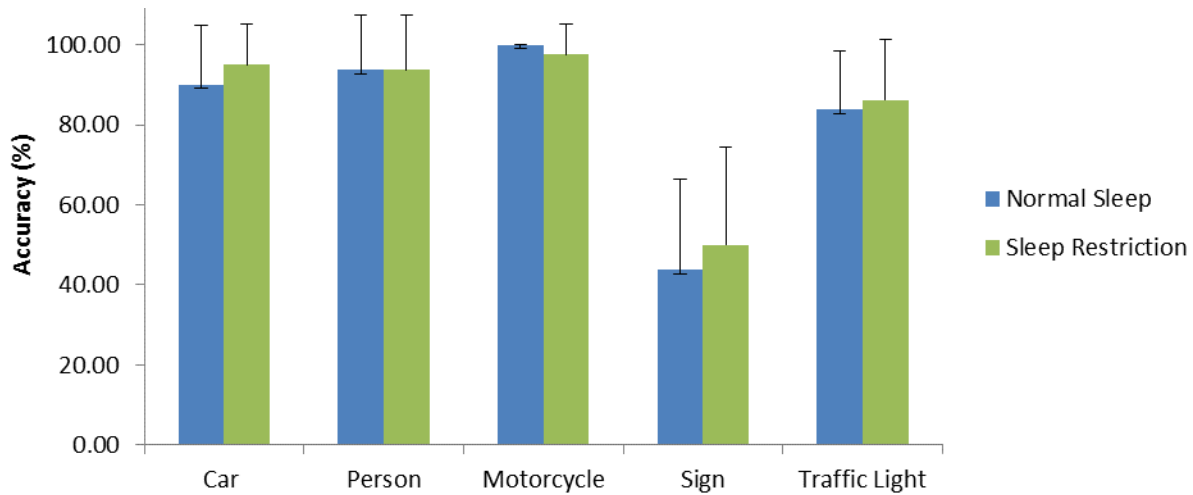


Figure 4-2. Change detection accuracy in urban driving scenes, by change target and sleep condition. Error bars represent standard deviation.

#### 4.3.4.3. Urban vs. Rural Comparisons

Analyses directly comparing change detection performance in urban and rural environments included only trials containing changes related to cars, motorcycles and signs, as these were the three targets appeared in both environments.

There was a significant interaction between driving environment and change target,  $F(1.5, 28.8) = 43.16$ ,  $p < .001$ , partial  $\eta^2 = .69$ ,  $\epsilon = .76$ . Participants were less accurate at identify changes to signs in urban scenes compared with rural scenes. This is reflected in a significant main effect of environment,  $F(1, 19) = 79.09$ ,  $p < .001$ , partial  $\eta^2 = .81$ , and change target,  $F(1.2, 22.7) = 55.89$ ,  $p < .001$ , partial  $\eta^2 = .75$ ,  $\epsilon = .60$ , on accuracy. All pairwise comparisons were significant identifying that accuracy changes for each target were significantly different from all other targets [ $p < .01$ ]. There was no significant main effect of sleep on accuracy,  $F(1, 19) = 0.00$ ,  $p = 1.000$ , partial  $\eta^2 = .00$ . There was no significant interaction between environment and sleep,  $F(1, 19) = 3.31$ ,  $p = .085$ , partial  $\eta^2 = .15$ . Results are presented in Figure 4-3.

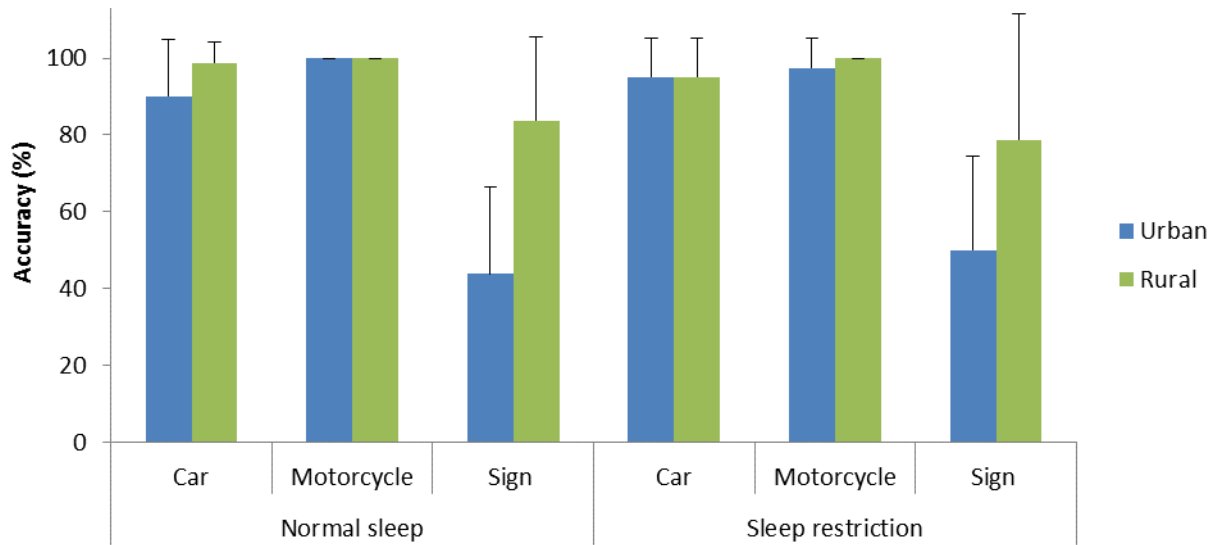


Figure 4-3. Mean accuracy for Change-present trials for each sleep condition separated by driving environment. Error bars represent standard deviation.

#### 4.3.5. Change Detection: Response Time

##### 4.3.5.1. Rural Scenes

RT analysis was undertaken for correct trials only. The tree condition was excluded from RT analysis due to the very low accuracy (only five participants detected any changes involving trees). Results are for 18 participants who had accurate trials for all four change targets. Sleep restriction did not significantly affect RT,  $F(1,17) = 0.59$ ,  $p = .452$ , partial  $\eta^2 = .03$ . However, there was a significant main effect of target type on RT,  $F(3,51) = 10.36$ ,  $p < .001$ , partial  $\eta^2 = .38$ . Pairwise comparison identified that RT to cars was significantly faster than to animals [ $p = .010$ ] or signs [ $p = .004$ ]. RT to motorcycle changes was also significantly faster than to animals [ $p = .022$ ] or signs [ $p = .006$ ]. There was no significant interaction between sleep condition and change target,  $F(1.4,5.3) = 0.77$ ,  $p = .457$ , partial  $\eta^2 = .16$ ,  $\epsilon = .33$ . Mean RTs are shown in Figure 4-4.

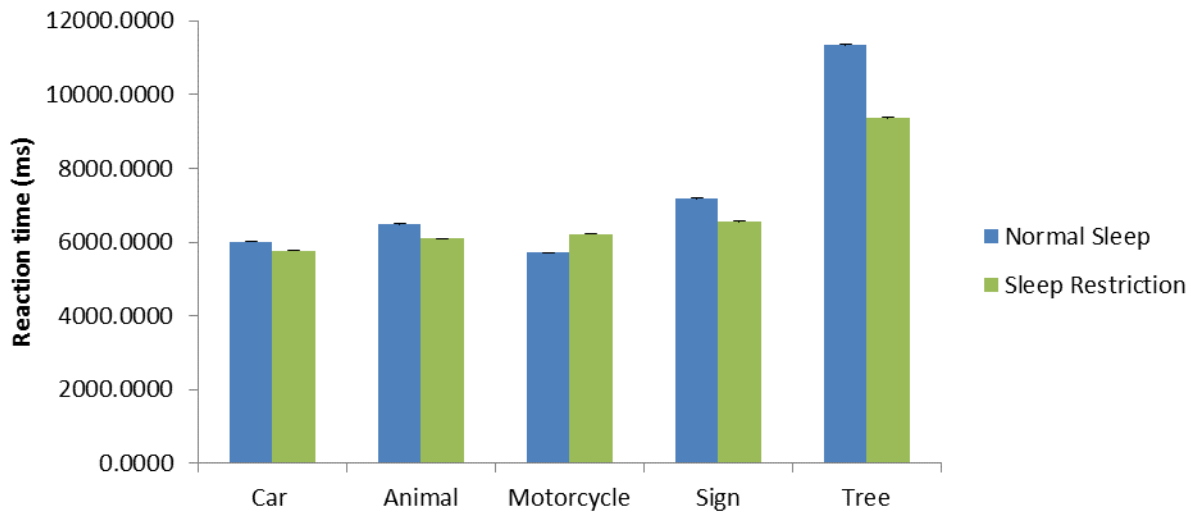


Figure 4-4. Mean RT for accurate change-present trials per target change target in rural scenes. Results are presented for both sleep conditions. The sample include 5 participants for tree changes and 18 participants for all other target types. Error bars represent standard deviation.

#### 4.3.5.2. Urban scenes

Statistical comparison was carried out on data from 17 participant with accurate trials for all five urban change targets. There was a significant main effect of sleep condition on RT,  $F(1,16) = 6.12$ ,  $p = .025$ , partial  $\eta^2 = .28$ , with participants requiring longer to correctly detect changes when they had experienced sleep restriction compared with normal sleep. There was also a significant main effect of change target RT,  $F(3.2,50.5) = 5.39$ ,  $p = .002$ , partial  $\eta^2 = .25$ ,  $\epsilon = 0.79$ . Pairwise comparison identified that RT to motorcycles was significantly faster than to people [ $p = .025$ ]. Similarly, reaction to motorcycles was significantly faster than to signs [ $p = .028$ ]. There was no significant interaction between sleep condition and change target,  $F(3.0,47.1) = 0.59$ ,  $p = .619$ , partial  $\eta^2 = .36$ ,  $\epsilon = 0.74$ . Results are presented in Figure 4-5.

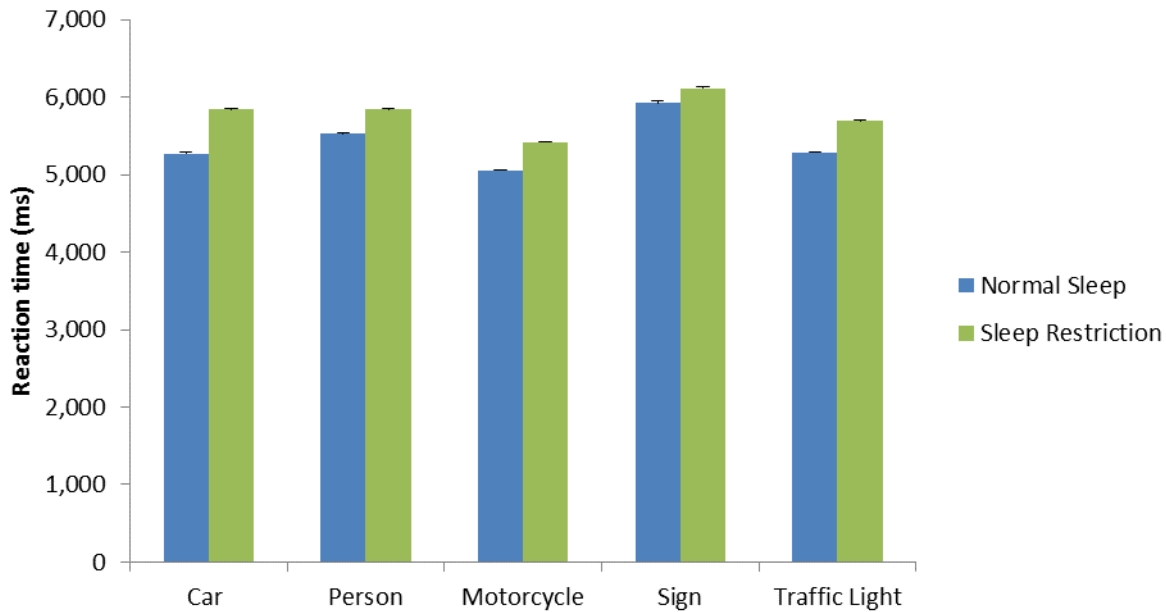


Figure 4-5. Mean RT for accurate change-present trials per change target type in urban scenes. Results are presented for both sleep conditions. Bars represent mean RT for 17. Error bars represent standard deviation.

#### 4.3.5.3. Urban vs. Rural Comparisons

Statistical comparison was carried out on data from 16 participants with accurate trials for the three target objects in both driving environments. There was a significant main effect of change target on RT,  $F(1.5, 21.9) = 13.27$ ,  $p < .001$ , partial  $\eta^2 = .47$ ,  $\epsilon = .73$ . Pairwise comparison identified that RTs for signs were significantly slower than for changes to either cars [ $p = .003$ ] or motorcycles [ $p = .005$ ]. There was no significant main effect of environment,  $F(1, 15) = 1.11$ ,  $p = .309$ , partial  $\eta^2 = .07$ , or sleep,  $F(1, 15) = 3.92$ ,  $p = .066$ , partial  $\eta^2 = .21$ . Results are presented in Figure 4-6.



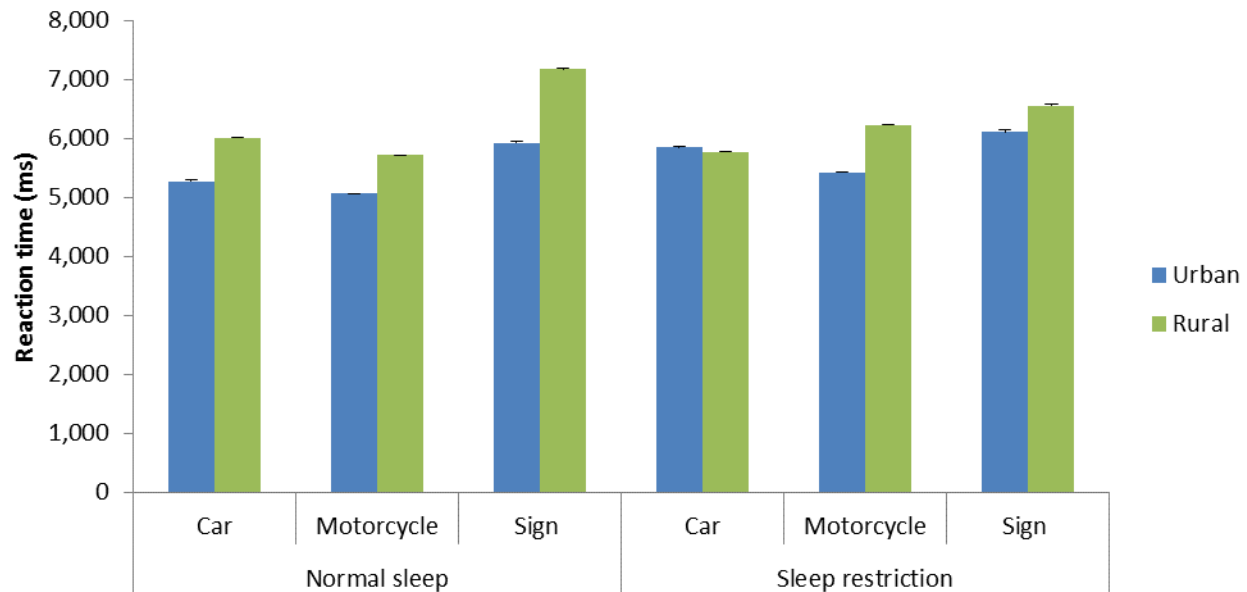


Figure 4-6: Mean RTs for each target type in each driving environment separated by sleep condition and environment ( $n = 16$ ). Error bars represent standard deviation.

#### 4.3.6. Change-absent Trials (Correct Rejections)

##### 4.3.6.1. Accuracy: Change-absent Trials

There was no significant main effect of environment [ $F(1,19) = 0.00$ ,  $p = 1.000$ , partial  $\eta^2 = .00$ , or sleep condition,  $F(1,19) = 1.21$ ,  $p = .285$ , partial  $\eta^2 = .06$ , on accuracy for correctly identifying no change had occurred. There was no significant interaction between environment and sleep condition on accuracy,  $F(1, 9) = 1.15$ ,  $p = .297$ , partial  $\eta^2 = .06$ .

##### 4.3.6.2. Response Time: Change-absent Trials

There was a significant interaction between environment and sleep on RT,  $F(1,19) = 6.81$ ,  $p = .017$ , partial  $\eta^2 = .26$ . Sleep restriction slowed RT in the rural environment but increased RT in the urban environment. There was a significant main effect of environment,  $F(1,19) = 9.08$ ,  $p = .007$ , partial  $\eta^2 = .32$ . Reactions were slower in urban ( $M = 8839$  ms,  $SEM = 724$ ) than in the rural ( $M = 8361$  ms,  $SEM = 732$ ) environment. There was no significant main effect of sleep,  $F(1,19) = 0.78$ ,  $p = .783$ , partial  $\eta^2 = .004$ . Results are presented in Figure 4-7.

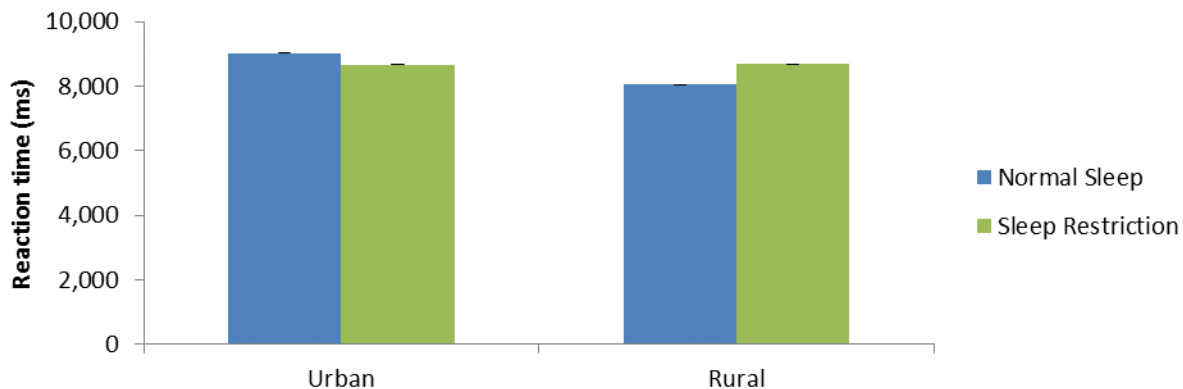


Figure 4-7. Mean RT for change-absent trials for each driving environment for both sleep conditions. Error bars represent standard deviation.

#### 4.3.7. Eye Movements: Fixations on Change Target

##### 4.3.7.1. Rural Scenes

For trials in which the target was correctly identified, eye movements to the change target were analysed. Two aspects of eye movement behaviour were analysed: the number of times that the change target was fixated and the total dwell time on the target, during each trial. Statistical analysis was conducted for 18 participants who had eye tracking measures for four of the five change targets. Data for trees was excluded from analysis due to the small number of participants with accurate trials. Mean values are presented in Table 4-3.

Due to small number of participants having accurate trials for trees, analysis was completed without trees for the 18 participants that fixated on each of the other targets. For the number of fixations there was no main effect of sleep,  $F(1,17) = 2.53$ ,  $p = .130$ , partial  $\eta^2 = .13$ . There was a significant main effect of change target,  $F(2.1,35.3) = 11.33$ ,  $p < .001$ , partial  $\eta^2 = .40$ ,  $\epsilon = .69$ . Pairwise comparison identified that signs were fixated on significantly more times than for any other change target [ $p \leq .02$ ]. There was a significant sleep by change target interaction,  $F(2.0,34.1) = 4.50$ ,  $p = .018$ , partial  $\eta^2 = .21$ ,  $\epsilon = .67$ . Following sleep restriction participants reduced the number of fixations on cars and animals but increased the number of fixations on motorcycles and signs.

For the total dwell time there was no main effect of sleep,  $F(1,17) = 0.02$ ,  $p = .889$ , partial  $\eta^2 = .00$ . There was a significant main effect of change target,  $F(2.5,42.1) = 8.08$ ,  $p < .001$ , partial  $\eta^2 = .32$ ,  $\epsilon = .83$ . Pairwise comparison identified that the dwell time on signs was significantly longer than for animals and motorcycles [ $p < .01$ ], but not significantly different to dwell time on cars. This suggests that participants needed to look at a sign for longer to process a change than for animals and motorcycles. There was a significant sleep by change target interaction,  $F(1,17) = 11.99$ ,  $p < .001$ , partial  $\eta^2 = .41$ . Following sleep restriction participants reduced the dwell time on cars and animals but increased the number dwell time on motorcycles and signs. This suggests that when sleepy participants need to look at motorcycles and signs for longer to be able to process that a change has occurred.

Table 4-3

*Average number of fixations (count) and dwell time (in milliseconds) on target, by target type and sleep condition for rural scenes*

		Normal Sleep <i>M (SD)</i>	Sleep restriction <i>M (SD)</i>
<b>Number of fixations on the target</b>	Car ( <i>n</i> = 20)	1.0 (0.5)	0.7 (0.6)
	Animal ( <i>n</i> = 20)	1.0 (0.5)	0.8 (0.9)
	Motorcycle ( <i>n</i> = 20)	0.7 (0.5)	0.9 (0.8)
	Sign ( <i>n</i> = 18)	1.1 (0.8)	1.8 (1.1)
	Tree ( <i>n</i> = 5)	5.1 (3.5)	3.4 (2.9)
<b>Total dwell time on the target (ms)</b>	Car ( <i>n</i> = 20)	402 (216)	324 (300)
	Animal ( <i>n</i> = 20)	353 (246)	366 (276)
	Motorcycle ( <i>n</i> = 20)	261 (215)	451 (524)
	Sign ( <i>n</i> = 18)	617 (470)	764 (485)
	Tree ( <i>n</i> = 5)	1784 (1377)	1553 (1449)

#### 4.3.7.2. Urban Scenes

As with the rural scenes, number of fixations and total dwell time on the change target were analysed for urban scenes in which the participant correctly identified the change. Mean values are presented in Table 4-4. Statistical analysis was conducted on results from 16 participants who had valid eye-tracking results for change-present trials.

For the number of fixations there was no main effect of sleep,  $F(1,15) = 1.77$ ,  $p = .203$ , partial  $\eta^2 = .11$ . There was a significant main effect of change target,  $F(4,60) = 21.85$ ,  $p < .001$ , partial  $\eta^2 = .59$ . Pairwise comparison identified that cars were fixated on significantly more times than for any other change target [ $p < .01$ ]. There was no significant sleep by change target interaction,  $F(3.3,49.6) = 0.84$ ,  $p = .491$ , partial  $\eta^2 = .05$ ,  $\epsilon = .83$ .

For the total dwell time there was no main effect of sleep,  $F(1,16) = 1.92$ ,  $p = .185$ , partial  $\eta^2 = .11$ . There was a significant main effect of change target,  $F(4,64) = 14.55$ ,  $p < .001$ , partial  $\eta^2 = .48$ . Pairwise comparison identified that dwell time on cars was significantly greater than for any other change target [ $p < .015$ ]. This suggests that participants needed to look at a cars for longer to process a change than for other change targets. There was no significant sleep by change target interaction,  $F(2.8,44.8) = 0.15$ ,  $p = .918$ , partial  $\eta^2 = .01$ .

Table 4-4

*Average number of fixations (count) and dwell time (in milliseconds) on target, by target type and sleep condition for urban scenes*

	Change Target	Normal Sleep <i>M (SD)</i>	Sleep Restriction <i>M (SD)</i>
<b>Number of fixations on the target</b>	Car ( <i>n</i> = 20)	1.6 (0.7)	1.8 (1.0)
	Person ( <i>n</i> = 19)	0.8 (0.5)	0.9 (0.7)
	Motorcycle ( <i>n</i> = 20)	1.0 (0.6)	0.9 (0.6)
	Sign ( <i>n</i> = 17)	0.7 (0.6)	1.2 (0.6)
	Traffic light ( <i>n</i> = 20)	1.0 (0.7)	0.9 (0.6)
<b>Total dwell time on the target (ms)</b>	Car ( <i>n</i> = 20)	619 (315)	746 (461)
	Person ( <i>n</i> = 20)	325 (241)	350 (311)
	Motorcycle ( <i>n</i> = 20)	383 (222)	424 (388)
	Sign ( <i>n</i> = 17)	480 (433)	567 (309)
	Traffic light ( <i>n</i> = 20)	393 (232)	426 (311)

#### 4.3.7.3. Urban vs Rural Comparison

The number of fixations made and total dwell time on car, motorcycle, and sign targets were compared between urban and rural driving environments. Analysis is only for those trials where changes were accurately identified. Results for 16 participants are presented in Tables 4-5 and 4-6.

Table 4-5

*Mean (and SD) number of fixations on each target type by driving environment.*

		Urban	Rural
<b>Normal Sleep</b>	Car	1.7 (0.7)	1.0 (0.5)
	Motorcycle	1.0 (0.6)	0.7 (0.5)
	Sign	0.7 (0.6)	1.2 (0.8)
<b>Sleep Restriction</b>	Car	1.9 (1.0)	0.8 (0.6)
	Motorcycle	1.0 (0.6)	1.0 (0.9)
	Sign	1.2 (0.8)	1.8 (1.1)

For the number of fixations there was a significant main effect of sleep,  $F(1,15) = 5.20$ ,  $p = .038$ , partial  $\eta^2 = .26$ . Participants made more fixations on the target following sleep restriction; however, this finding should be interpreted with caution given that the separate analyses in the urban and rural environments did not find a significant main effect of sleep. There was no significant effect of environment,

$F(1,15) = 1.75, p = .206$ , partial  $\eta^2 = .10$ . There was a significant main effect of change target,  $F(2,30) = 7.28, p = .003$ , partial  $\eta^2 = .33$ . Pairwise comparison identified that vehicles were fixated on significantly more times than motorcycles [ $p < .01$ ]. There was no significant sleep by environment interaction,  $F(1,15) = 0.002, p = .962$ , partial  $\eta^2 = .00$ . There was no significant sleep by change target interaction,  $F(1.4,20.6) = 2.56, p = .116$ , partial  $\eta^2 = .15$ . There was a significant environment by target interaction,  $F(1.5,22.6) = 44.29, p < .001$ , partial  $\eta^2 = .75$ . In urban environments the number of fixations on motorcycle and car targets increased while the number of fixations on sign targets decreased compared to rural. The three-way sleep, environment, change target interaction was not significant,  $F(2,30) = 2.21, p = .128$ , partial  $\eta^2 = .13$ .

Table 4-6

*Mean (and SD) of dwell time on each target type by driving environment, in milliseconds.*

		Urban	Rural
Normal Sleep	Car	662 (325)	408 (237)
	Motorcycle	407 (232)	278 (226)
	Sign	469 (445)	678 (460)
Sleep Restriction	Car	807 (470)	348 (330)
	Motorcycle	470 (415)	508 (572)
	Sign	586 (308)	812 (492)

For the dwell time there was no a significant effect of sleep,  $F(1,15) = 3.98, p = .064$ , partial  $\eta^2 = .21$ , or driving environment,  $F(1,15) = 0.96, p = .342$ , partial  $\eta^2 = .06$ . There was a significant main effect of change target,  $F(2,30) = 8.97, p = .001$ , partial  $\eta^2 = .37$ . Pairwise comparison identified that dwell time was significantly greater for motorcycles than for either signs [ $p = .012$ ] or cars [ $p = .006$ ]. This suggests that it took longer to process changes involving motorcycles. There was no significant interaction between sleep and driving environment,  $F(1,15) = 0.01, p = .940$ , partial  $\eta^2 = .00$ . There was no significant sleep by change target interaction,  $F(1.5,22.1) = 0.60, p = .508$ , partial  $\eta^2 = .04$ .

There was a significant environment by change target interaction,  $F(2,30) = 23.15, p < .001$ , partial  $\eta^2 = .61$ . In urban environments dwell time increased for motorcycles and cars compared to in rural environment. This suggests that participants needed to fixate on the target for longer to process the change. The opposite was found for signs where participants spent longer looking at the target in the rural environment before identifying the change than in urban. The Sleep, Environment, change target interaction was not significant,  $F(2,30) = 1.91, p = .165$ , partial  $\eta^2 = .17$ .

#### 4.3.8. Eye Movement: Non-Target Fixation Patterns

To examine the effect of sleep loss on scanning patterns more generally, the overall number of fixations made in each trial and the average duration of fixations (regardless of which interest area they were in) was compared between NS and SR conditions. As shown in Table 4-7, sleep restriction was not associated with systematic differences in the number or duration of fixations made.

Two other measures that may be associated with sleepiness are pupil size (which has been used as an indicator of workload) and visual tunnelling. To measure visual tunnelling, fixation location coordinates were averaged for each participant to calculate a central gaze point (i.e., that individual's central fixation point). The 50<sup>th</sup>, 70<sup>th</sup> and 90<sup>th</sup> percentile distances from this central fixation point were then compared between NS and SR conditions. If tunnelling occurred, then the average fixation distance from the central point should be smaller. As shown in Table 4-7, neither pupil size nor any of the visual tunnelling measures were significantly different between sleep conditions.

Table 4-7

*Mean (and SD) for general ocular measures for both sleep conditions.*

	Normal Sleep	Sleep Restriction	Comparison
Average number of fixations per trial	16.6 (5.5)	16.7 (5.5)	$t(19) = 0.10, p = .924$
Average fixation duration (ms)	327 (40)	328 (41)	$t(19) = 0.26, p = .796$
Average pupil size	1145 (324)	1130 (285)	$t(19) = 0.40, p = .696$
50 <sup>th</sup> percentile deviation from average fixed location (pixels)	204 (24)	206 (20)	$t(19) = 0.41, p = .686$
70 <sup>th</sup> percentile deviation from average fixed location (pixels)	291 (34)	293 (32)	$t(19) = 0.42, p = .677$
90 <sup>th</sup> percentile deviation from average fixed location (pixels)	462 (46)	462 (40)	$t(19) = 0.07, p = .944$

Finally, in order to compare overall scanning patterns, we compared proportion of time participants spent fixating on five key interest areas (road, left of road, right of road, horizon, and sky) between urban and rural scenes, and NS and SR conditions. This analysis included data from correct change-absent trials, as in Experiment 1. The mean percent of trial time spent scanning each region of the images is shown in Table 4-8.

In terms of general scanning pattern there was no a significant effect of sleep,  $F(1,19) = 1.94, p = .180$ , partial  $\eta^2 = .09$ . There was a significant effect of environment,  $F(1,19) = 66.01, p < .001$ , partial  $\eta^2 = .78$ . There was a significant main effect of interest area,  $F(1.8,34.6) = 49.24, p < .001$ , partial  $\eta^2 = .72$ . Pairwise comparison identified that the proportion of dwell time in sky was significantly less than for any other area [ $p < .001$ ]. Dwell time to the left of the road occurred for a significantly greater proportion of time than to the right of the road [ $p < .001$ ], and a significantly greater proportion of dwell time was spent looking in the horizon than to the right of the road [ $p < .001$ ]. There was no significant sleep by environment interaction,  $F(1,19) = 0.68, p = .420$ , partial  $\eta^2 = .04$ . Nor was the interaction between sleep

and interest area significant,  $F(2.0,37.9) = 2.10$ ,  $p = .137$ , partial  $\eta^2 = .10$ . There was a significant environment by interest area interaction,  $F(1.8,34.1) = 7.05$ ,  $p = .004$ , partial  $\eta^2 = .27$ . In urban environments proportion of dwell time on the left of road, sky and horizon was increased compared to rural environments. In contrast the proportion of dwell time spent on the road was reduced in rural compared to urban environments. Dwell time on the right of the road was similar between environments. The sleep, environment, interest area interaction was not significant,  $F(2.0,37.3) = 0.27$ ,  $p = .765$ , partial  $\eta^2 = .01$ .

Table 4-8

*Mean (and SD) dwell time in each interest area during correct Change-Absent trials by driving environment and sleep condition.*

	Interest Area	Urban	Rural
<b>Normal Sleep</b>	% dwell time on the road	30 (10)	34 (12)
	% dwell time left of road	32 (6)	31 (5)
	% dwell time right of the road	23 (6)	23 (5)
	% dwell time on the sky	16 (4)	9 (3)
	% dwell time on the horizon	32 (7)	26 (7)
<b>Sleep restriction</b>	% dwell time on the road	28 (8)	30 (13)
	% dwell time left of road	32 (6)	31 (13)
	% dwell time right of the road	25 (6)	25 (7)
	% dwell time on the sky	15 (5)	10 (3)
	% dwell time on the horizon	34 (6)	28 (6)

#### 4.4. Discussion

The aim of the current study was to examine the impact of sleep loss on drivers' change detection ability when viewing urban and rural driving scenes. All participants were experienced drivers who drove at least weekly and were familiar with the locations depicted in the stimulus images. The results revealed that sleep loss has minimal impact on change detection performance, although this varied as a function of driving environment. Consistent with the results of Experiment 1, the type of object that changed was the most consistent impacting factor on change detection.

##### 4.4.1. Effect of Sleep Loss

Overall, mild sleep loss does not impact drivers' ability to accurately identify changes to either rural or urban driving scenes. Participants were compliant with protocol, restricting their sleep to 5 hours on one occasion. This resulted in them feeling significantly sleepier. The level of sleepiness experienced did not impair accuracy at the change detection task. However, sleep loss significantly slowed RT for detecting changes in urban scenes, but not rural scenes. On average, the time required to correctly identify



changes in urban environments was 370 milliseconds longer. While this may seem trivial, at a travel speed of 60 km/h this equates to an extra 6.2 metres travelled, which could be the difference between a collision and a near-miss. Sleep-related crashes on rural roads are often out of lane events, most probably caused by the driver falling asleep or having a micro sleep (Horne & Reyner, 1995). One possibility is that the underlying characteristics of sleep-related crashes differ between urban and rural roads. In light of the current findings it is possible that sleep-related crashes in more complex urban environments may be influenced by a slowed RT to change detection rather than explicitly falling asleep. The current results are particularly interesting because the slowed RT was apparent in urban but not rural scenes. Urban scenes are more visually cluttered, presenting more items within an environment to which participants needed to attend. As information processing and planning are skills known to be impaired by sleep loss (Horne, 2012), it might be expected that change detection is more influenced by sleep loss in situations where a greater amount of information is needed to be processed.

A key component of the process to identify changes is the ability to visually scan a scene. Sleepiness is known to impact oculomotor function (De Gennaro et al., 2000; Fransson et al., 2008), which impairs visual search performance (De Gennaro et al., 2001). One reported outcome of sleepiness is that drivers “tunnel” their vision and spend a greater amount of gaze time looking at the centre of the road (Fors et al., 2013). It has been reported that participants who are sleepy are more likely to maintain their gaze within the central area of the road and reduce attention to the periphery. If tunnelling were to occur it would be expected that a greater proportion of sleep restriction fixations would have occurred within a smaller location (number of pixels). However, fixations at the 50<sup>th</sup>, 70<sup>th</sup> and 90<sup>th</sup> percentile were similar regardless of whether participants were alert or sleepy suggesting the visual tunnelling did not occur.

Although accuracy for change-absent trials was consistently near perfect (>98%) regardless of driving environment or sleep condition, there was a significant sleep by environment interaction on RT. Previous research has consistently reported accuracy for change-absent trials to be greater than for change-present trials (e.g., Josephs et al., 2016). However, the impact of physiological stresses such as sleepiness have on change-absent trials have not been well researched. Following sleep loss RTs reduced for change-absent urban images but increased for rural images. Within each environment the change between sleep conditions was not significant. The interaction effect appears to be influenced by the RT in rural environment being significantly faster than the RT for urban changes under the normal sleep condition. In contrast, when sleep was restricted the RT advantage for rural scenes is no longer apparent. This suggests that it is possible for alert participants to quickly able to scan the simple rural environments and accurately determine that no change has occurred, but feeling sleepy impairs this rapid scanning ability. In contrast, complex urban environments require longer to scan, regardless of the participant’s level of alertness. This is possibly influenced by differing visual search strategies between the two environments.

Overall, participants spent a greater proportion of time looking at the road area in the rural scenes, compared with urban scenes, which is consistent with the results of Experiment 1. Additionally, the dwell time required to correctly identify changes involving cars and motorcycles was significantly longer in urban environments, compared with rural environments, which indicates that participants had more

difficulty identifying these changes in urban environments. This suggests that the visual complexity of the scene has a strong influence on how the same targets are visually processed.

Sleep loss also produced differential effects on visual dwell time for different change targets in rural scenes. Participants showed shorter dwell times for cars and animals, but longer dwell times for motorcycles and signs, when they were sleepy compared to when they had a normal night's sleep. This suggests that in order to achieve the same accuracy participants needed to look more intently and took longer to process motorcycle and sign changes when sleepy. However, fewer participants were accurate at identifying sign changes. This finding may be associated with the potential threat associated with the change target. Under the pressure of sleep loss participants maintained that ability to rapidly identify those hazards with greatest potential for immediate danger – namely, cars and kangaroos. In contrast, a subtle impairment is apparent to participant ability to perceive the less hazardous environment elements (motorcycles and signs). An additional influencing factor is the possibility that those participants who could accurately identify changes to signs were those who had greater aptitude for the task, and these high performing participants were less affected by sleep loss.

#### ***4.4.2. Effect of Target Type***

Of all the factors considered in Experiment 2, the type of change target demonstrated the largest and most consistent effects on change detection accuracy and RT. This was consistently the case for both rural and urban scenes, and is consistent with the results of Experiment 1. The greater accuracy and speed for identifying changes to other road users compared to trees might be expected as previous research has reported greater accuracy for driving-relevant compared to irrelevant targets (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002). However, previous examples of task irrelevant change targets were all static objects (e.g., mailboxes). The current work furthers these findings by reporting that accuracy for animals, a dynamic task irrelevant change target, was also lower than accuracy for detecting changes to task relevant road users.

The amount of time participants spend looking at an accurately identified change target gives an indication of how long it takes to cognitively process the fact that a change has occurred. Participants made the greatest number of fixations and spent the longest amount of dwell time looking at signs compared to other types of targets (excluding trees, which most participants failed to notice altogether). Previous research has demonstrated that participants often fail to notice changes to familiar signs (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007); in addition to replicating this finding, the current study's eye-tracking results suggests that changes to signs are quite demanding to process.

Interestingly, for rural images dwell time on signs was not significantly longer than for cars, despite RTs being significantly longer for signs. This difference could be explained by differing visual search strategies applied to the two target types. It is possible that participants actively look for changes to cars before they look for changes to signs despite needing a similar amount of processing time for both. Similarly in urban images, cars were fixated on more often and for longer than any other change target. As all participants were drivers and had been instructed that the images being viewed were from a driver's perspective, it is possible that other cars were considered the most task relevant change target,

further adding to the evidence that task relevance is highly influential on ability to detect change (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002).

#### ***4.4.3. Effect of Driving Environment: Urban vs Rural***

Participants were more accurate at identifying changes in the less complex rural scenes, compared with urban scenes. This is consistent with results from Experiment 1. However, in contrast to Experiment 1, the effect of driving environment on RT was not significant. This difference may be an artefact of experimental design. In Experiment 2 images were presented in blocks relevant to the environment, therefore all urban images were presented in succession. In contrast, Experiment 1 presented rural and urban images in an intermixed order. It is possible that in Experiment 2 participants became adjusted to the consistent viewing of urban images and adopted a more efficient scanning strategy, thereby maintaining RT.

When considering urban and rural images together, the effect of sleep loss on RT approached significance. This did not support the finding of RT impairment for urban images alone. In part this is likely to be due to the reduced number of change targets being considered (i.e., because the urban-rural direct comparison included only the three types of targets that changed in both environments, namely cars, motorcycles and signs) as well as the conflicting influence of sleep loss for the different environments.

#### ***4.4.4. Summary***

Overall the results of Experiment 2 indicate that sleep loss has a slight impact on change detection efficiency but not accuracy, in urban environments. Target type has a large impact in drivers' ability to detect changes, specifically, drivers are more efficient at detecting changes to other road users than static objects (trees and signs). As in Experiment 1 drivers are better at detecting changes in rural scenes compared to urban scenes, which is likely because there is less visual clutter in rural areas.

The slowed RT to urban images under sleep restriction is concerning as it may be a contributing factor towards crashes. A limitation of Experiments 1 and 2 is that RT was measured in response to a static image, in contrast when driving a person moves through an environment and has added task load relating to vehicle operation. These added pressures may be sufficient to exacerbate any impaired change detection ability due to sleepiness. This is also important in relation to previous research which reports change blindness to be more common for participants engaged in a secondary task (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002). Experiment 3 will consider response to dynamic changes and observe whether urban/rural differences remain under the added pressure of a dual task paradigm whereby participants will be required to drive as well as detect changes.

A further limitation of the current work is that participants were directed to look for changes in each image pair. In reality drivers are not cued to the potential presence of a change. Experiment 3 will address this by considering detection of both expected and unexpected changes.

## 5. Experiment 3

### 5.1. Background and Rationale

Experimental research is vital for understanding the road safety implications of driving when sleepy in part because crash data underestimates the number of sleep-related crashes (Åkerstedt, 2000), providing only crude understanding of driver sleepiness with limited application for developing interventions. Unlike alcohol, sleepiness cannot be measured by a simple index such as blood-alcohol concentration (BAC). Sleepiness is particularly difficult to retrospectively quantify in real-life driving situations (Radun et al., 2013). Driving simulators offer a safe, controllable and measurable environment for experimental investigation of driver sleepiness (Liu et al., 2009).

Driving simulator studies investigating sleepiness commonly replicate the most frequent environmental surrounding of on-road sleep related crashes. It is necessary to use repeated-measures designs because of large individual differences in the response to sleep loss (Van Dongen et al., 2004), so participants typically complete at least two simulated drives under varying levels of alertness. Studies are commonly conducted at night or during the circadian low of the afternoon (e.g., Anund et al., 2008; Filtiness et al., 2012; Horne & Reyner, 1996). The driving scenario is usually long (>1.5 hour), represents a monotonous road and provides limited interaction with other road users. The presented road is often a motorway or rural highway. Predominantly straight roads are most often used with slight curves at spaced intervals. The curves are introduced to ensure that participants must actively control the vehicle in order to stay on the road. These conditions are selected to enhance the potential for sleepiness and reduce the chance the experimental protocol influencing alertness. This approach may be considered as presenting the “worst case scenario” under which to test participant’s performance. Using such paradigms it has been consistently demonstrated that driver sleepiness increases the number and frequency of out of lane events (e.g., Horne & Reyner, 1996), variability in lane positioning (e.g. Anund et al., 2008; Forsman et al., 2013) and variability in speed control (e.g., Matthews et al., 2012). These investigations (and many others) have provided important insight into driver sleepiness, but in each case the influence of sleepiness on driving performance were examined in similar high speed monotonous conditions.

Results from Experiment 2 provide some suggestion that sleep loss may impair change detection in urban environments. However, the experimental protocol used static images in a “flicker” change detection paradigm. The extent to which these findings generalise to real-world driving is unclear, given that physically driving requires detection of dynamic change targets while under the added task load of maintaining vehicle control. It is important to investigate change blindness under realistic driving conditions as previous research has demonstrated change blindness to be stronger in dynamic stimuli compared to static stimuli (Velichkovsky et al., 2002). This suggests that change blindness may be more likely to occur during simulated driving (and potentially real driving) than in lab-based experiments.

Change blindness is particularly likely to occur when visual changes take place during a disruption to the visual scene, such as when a person is blinking, making an eye movement, or has their view obscured briefly (e.g., McConkie & Currie, 1996; Pashler, 1988; Rensink et al., 1997; Simons & Levin, 1998), as the disruption masks visual transients that would otherwise make the change obvious to the observer. Previous research has indicated that there are no functional differences in change blindness, regardless

of whether it results from a natural eye movement or an imposed interruption (Velichkovsky et al., 2002), which means it is possible to realistically simulate the type of change blindness that can occur during eye-blinks by inserting visual occlusions within a driving simulator scenario. Several studies using simulators to investigate change blindness have masked changes with brief occlusion periods (Lee et al., 2007; Shinoda et al., 2001; Velichkovsky et al., 2002; White & Caird, 2010), similar to the blank screens used in flicker tasks, whereas others have changes occur more naturalistically, for example changing a sign between repeated drives on the same road (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007). The current work uses a combination of the two approaches in order to consider the impact of sleep loss on change blindness for both cued changes, indicated by a blackout screen, and unexpected changes occurring out of sight but without a visual cue (blackout).

The current work presents a new approach to investigating driver sleepiness. Instead of targeting extreme out-of-lane events (e.g., following a micro sleep), the focus is on understanding subtle impairments associated with sleepiness and visual attention, specifically drivers' ability to detect expected and unexpected changes within their environment. Participants' sleep was restricted to a level expected to invoke sleepiness (5 hours instead of 8 hours), but not so extreme as to expect participants would fall asleep during the study. In addition to the overall aim of investigating the impact of sleep loss on change detection while driving, the study was also designed to assess whether there were systematic differences in change detection ability depending on the safety relevance of the change (high vs. low) and/or the driving environment (urban vs. rural) in which it occurred.

## 5.2. Method

### 5.2.1. Participants

Twenty-one drivers (12 female, 9 male) aged 18-30 years ( $M = 23.1$ ,  $SD = 3.9$ ) provided informed consent and voluntarily participated. Participants received AUD\$120 compensation for their time following completion of three study sessions, plus an allowance to cover travel expenses to attend the sleep restriction session. All participants had been driving unsupervised for at least one year, holding either a full ( $n = 13$ ) or second year provisional (P2;  $n = 7$ ) Australian licence. One participant held an overseas licence but had been driving regularly in Australia for more than one year and was sufficiently familiar with Australian road rules and conditions.

Participants were pre-screened to ensure they met relevant inclusion criteria. Specifically, participants were required to be non-smokers, regular drivers, and low consumers of caffeine (less than 5 times per day). It was also required that participants could not have a sleep disorder, work late-night shifts, or take daytime naps. Participants who experienced simulator motion sickness during practice drives did not continue with the study. Eleven participants were excluded after attending the initial session, eight due to simulator sickness, three chose to withdraw. Two additional participants completed the first two sessions but were unable to complete the third session.

One participant recorded only 5 hours sleep on the non-restricted night, however they remained in bed for an additional 2 hours and 20 minutes, possibly due to misunderstanding the task instructions.

Statistical analyses were conducted with this participant in the dataset, but significance levels were not affected by their exclusion so they were included in the final sample analysed.

Ethical approval (1500000653) was granted by Queensland University of Technology Human Research Ethics Committee prior to project commencement.

### **5.2.2. Apparatus**

The study was conducted using CARRS-Q's advanced driving simulator. This simulator is composed of a complete automatic Holden Commodore vehicle with working controls and instruments. The advanced driving simulator uses SCANeR™ studio software version 1.4 with eight computers, projectors and a six degree of freedom (6DOF) motion platform (Emotion 1500, REXROTH, Boxtel, Netherlands) that can move in three dimensions. When seated in the simulator vehicle, the driver is immersed in a virtual environment which includes a 180° forward field of view composed of three projector screens (with image input from three RGB video projectors), simulated rearview mirror images on LCD screens, surround sound for engine and environment noise, real car cabin, simulated vehicle motion and a steering wheel which provides force feedback (see Figure 5-1). The road and environment used conforms to current Australian Standards. The rendering capabilities of this simulator enable it to display a realistic driving environment (Canberra city, suburban and rural environments).



Photo by Sonja de Sterke

*Figure 5-1.* Photo of the advanced driving simulator showing size of main screens and motion platform.

### **5.2.3. Road Network**

Two matched driving scenarios were programmed (A and B), one for each of the study sessions. Each scenario required participants to drive 5 laps of an 11.3 km long circuit, with the whole drive taking approximately 45 minutes. An aerial view of part of the circuit, representing the civic area of Canberra (Northbourne Avenue and London Circuit) is shown in Figure 5-2. Approximately 50% of the driving time for each circuit was spent in an urban environment and 50% in a rural environment. The circuit was designed to be repetitive so that participants quickly became familiar with the driving route. All participants completed a practice drive (2 laps) to become familiar with the circuit, driving controls and procedure for change detection.



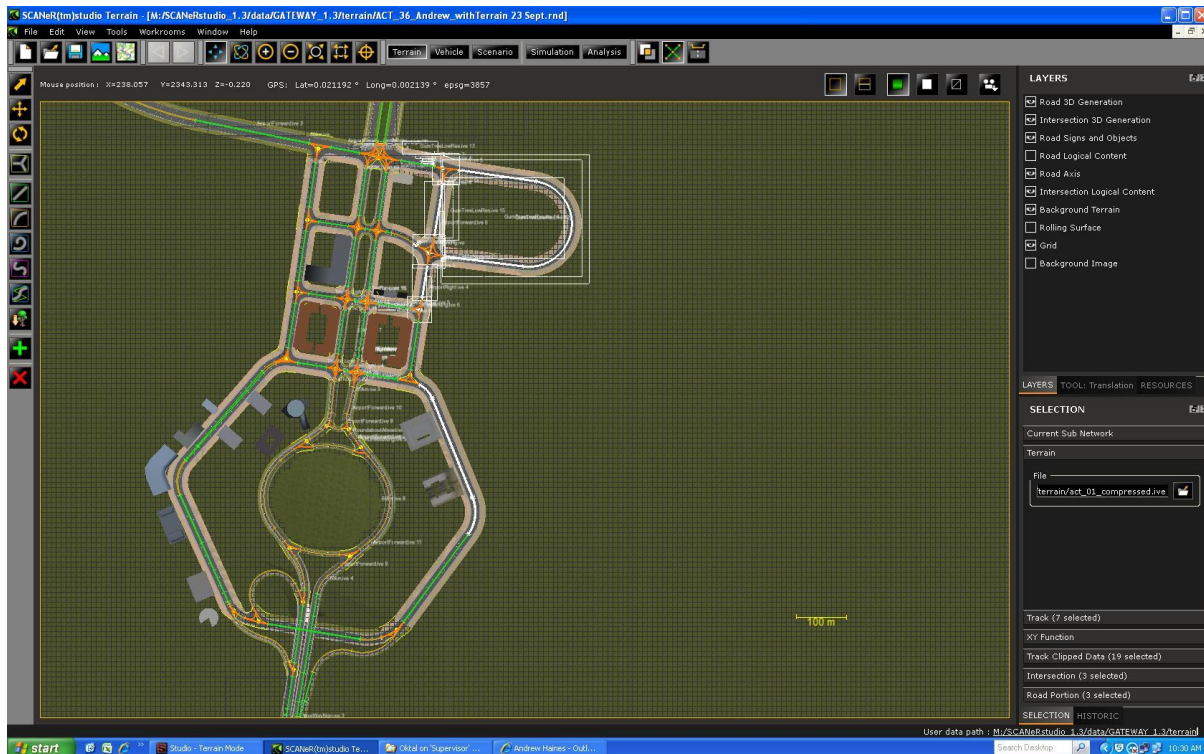


Figure 5-2. Screen shot of the civic area of Canberra within the simulator operating program.

The urban section of the circuit was based on aspects of Canberra including parts of civic, the inner north and the Parliamentary Triangle. The posted speed limit was 60 km/h.

The rural section of the circuit was a fictional road developed to be representative of roads in country NSW in the Canberra region. The rural road had one carriage way in each direction of travel and a hard shoulder on either side of the road. The road was predominantly straight with two gentle curve sections. During the rural section of the circuit there was one lead vehicle ahead of the participant, this was scripted to adjust its speed in relation to the participant, ensuring that it was always present but allowed participants to travel at a speed they were comfortable with. There were fewer vehicles present in the rural section, compared with the urban section, to maintain realism. The posted speed limit in the rural environment was 100 km/h.

## 5.2.4. Stimuli

### 5.2.4.1. Cued Change Detection

During each drive each participant experienced 20 cued change detection events (four per circuit lap). During a change detection event the simulator screens went black for 500 ms and then returned to either an identical scene or a scene with one difference. Blackouts were programmed to occur in response to a participant driving over a trigger point, which meant that the time between each blackout differed depending on the speed of the participant's driving.



Within each circuit lap there were two blackouts in the urban road section and two in the rural section. In each drive there were 12 change-present trials (6 urban, 6 rural) and 8 change-absent trials (4 urban, 4 rural). The practice drive contained 8 blackouts (3 change-absent), to allow participants to familiarise themselves with the task prior to the first experimental session. Following a blackout participants were asked to verbally report if they had noticed a change and if so to identify what the change was.

Among change-present trials the safety relevance of the change was manipulated, so that half of the changes in each environment had high safety relevance and half had low safety relevance, and the types of change events were matched in urban and rural environments. Table 5-1 lists examples of the changes implemented for each safety relevance level and driving environment.

Table 5-1

*Examples of high and low safety relevance changes following blackouts in Experiment 3*

	Low Safety Relevance	High Safety Relevance
Urban	<ul style="list-style-type: none"> <li>- Parked cars on roadside change colour</li> <li>- Advisory road sign changes to a different sign (one way sign to parking sign)</li> <li>- Cyclist travelling towards the participant on the opposite side of the road moves from the hard shoulder to the road.</li> </ul>	<ul style="list-style-type: none"> <li>- Car travelling towards participant moves into participant's lane (head on-collision)</li> <li>- Cyclist travelling ahead moves from hard shoulder to centre of participant's lane (rear-end collision)</li> <li>- Speed limit sign decreases by 10 km/h</li> </ul>
Rural	<ul style="list-style-type: none"> <li>- Parked cars on roadside change colour.</li> <li>- Advisory road sign changes to a different sign (yellow diamond sign to city ahead)</li> <li>- Tractor travelling towards the participant on the opposite side of the road moves from the hard shoulder to the road.</li> </ul>	<ul style="list-style-type: none"> <li>- Car travelling towards participant moves into participant's lane (head on-collision)</li> <li>- Tractor travelling ahead moves from hard shoulder to centre of participant's lane (rear-end collision)</li> <li>- Speed limit sign decreases by 10 km/h</li> </ul>

The same change events were used for each scenario but the order, timing and location of changes were varied to prevent participants anticipating changes in later stages of the study. Characteristics of the changed target (e.g., vehicle colour) were also varied between scenarios to further preclude participants anticipating upcoming changes. All objects used as change targets appeared multiple times throughout the drive in order to familiarise participants to their presence and ensure that participants did not habitually associate specific road users (e.g., a blue car) with any of the change detection events.

Change-present trials were considered "correct" if the observer correctly identified the change target, but were considered "incorrect" if they reported no change or incorrectly identified the change target. Change-absent trials were considered "correct" if the observer reported no change, and were considered "incorrect" if they indicated a change had occurred.

#### 5.2.4.2. Unexpected Change Detection

In addition to the changes following blackouts, four unexpected changes occurred during each drive. Detection of these change events was "uncued" in that there was no blackout screen to prompt participants to respond. One change occurred during each lap except lap 1. The lap associated with each

change varied between scenarios. No unexpected changes occurred during the practice drive. Three events were low safety relevance and one of high safety relevance.

The unexpected changes with low safety relevance involved:

- Non-essential road signs changed from English to German language. These signs were present at in three locations in the lap, and the change occurred for a single lap of the drive.
- Yellow diagonal hatching markings on the road changing to white square hatching. Hatchings were present at two intersections in each lap (one urban, one rural), and the change occurred for a single lap of the drive.
- The car immediately in front of the participant changed colour. This change occurred when the lead vehicle was briefly out of view, as the vehicles were travelling around a bend, but was timed so that it was not feasible that a different car could have actually joined the road ahead of the participant.

In the high safety relevance change, the participant was stopped at an urban signal-controlled intersection. Initially one pedestrian was crossing the road. As the pedestrian was crossing, a large vehicle passed and obscured the participant's view, and after the vehicle disappeared five pedestrians were visible.

Participants were instructed to verbalise anything that they noticed during the drive that was unusual or different to what they would expect. They were told they could use this response at any time, not exclusively following a blackout, to distinguish this from their responses following blackout periods.

Responses to unexpected changes were deemed correct if the participant identified the change correctly during the drive and incorrect if they did not mention the change at any time during the drive.

#### ***5.2.5. Self-Report Measures***

During the introductory session participants completed a demographics questionnaire. This included questions confirming participants met the screening criteria and description of their usual driving exposure and behaviour. In addition, the Epworth Sleepiness Scale (ESS; Johns, 1991) was used to identify if any participants experienced excessive day time sleepiness.

For the three nights prior to each experimental study session participants were required to keep daily sleep diaries recording their bed time, estimated sleep onset, night time wakings, and morning awakening and rising times.

Participants were asked report their subjective sleepiness on the KSS, as in Experiment 2. KSS ratings were given before the drive started, twice per lap (once in the rural and once in the urban section) and at the end of the drive. An average KSS rating was calculated to compare overall sleepiness in the normal sleep vs. sleep restriction sessions.

Sleep related eye symptoms (5 point scale) and effort to stay awake (7 point scale) were recorded post drive, and subjective workload (i.e., task demand) was measured post drive using the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) on a 10 cm visual analogue scale. For all post-drive measures, higher scores indicated higher levels of the items (i.e., increased sleepiness, effort, or workload).

### 5.2.6. Procedure

Participants attended three sessions in the advanced driving simulator, held on separate days at least three ( $M = 10.4$ ,  $SD = 10.0$ ) days apart. These comprised one introductory familiarisation session followed by two study sessions. Study sessions were conducted at either 10.30am (8 participants), 12.00 noon (2 participants), 1.30pm (5 participants) or 3.00pm (6 participants). Participants completed both experimental sessions at the same time of day. One study session was completed following a normal night of sleep and one after sleep restriction to five hours, with the order of sessions counterbalanced between participants.

During the introductory session participants provided written informed consent, completed background demographic questionnaires and practiced using the driving simulator. Participants were screened for motion sickness and simulator sickness. Those displaying any signs of sickness were not invited to complete the main study sessions. During the introductory session participants were given an Actiwatch and sleep diary with instructions on how to record their sleep leading up to the two study sessions.

In the two study sessions, participants provided their sleep diary and Actiwatch to a research assistant, who checked the data to ensure compliance with the required hours of sleep. Once this was confirmed, the participant was shown into the driving simulator and attempts were made to calibrate the eye-tracker calibrated.

Driving scenario order was counterbalanced both between sessions and sleep condition. Participants drove one scenario per session and the drive was completed without breaks.

### 5.2.7. Data Analysis

All statistical analyses were conducted using SPSS 21 statistical software. An alpha level of .05 was used to determine statistical significance. The experiment protocol was designed to include eye-tracking and vehicle control measures; however, due to technical difficulties this data was not available.

Paired  $t$ -tests were used to compare results between the Normal Sleep (NS) and Sleep Restriction (SR) conditions for subjective workload and sleepiness measures.

Accuracy for blackout change-present trials were analysed using RM-ANOVA with three within-subjects factors of Sleep Condition (2 levels: NS, SR), safety relevance (2 levels: high, low) and driving environment (2 levels: urban, rural).

McNemar's test was used to compare detection of unexpected changes between NS and SR conditions.

## 5.3. Results

### 5.3.1. Participants

Participants were frequent drivers ( $M = 6.5$  hours/week,  $SD = 3.4$ , range 2-14), covering a mean of 189 km per week ( $SD = 143$ , range 8-450). No participants would be considered to have excessive daytime sleepiness, as indicated by an ESS score  $>12$  ( $M = 5.4$ ,  $SD = 3.5$ , range 1-12). However, four participants scored between 10 and 12, indicating that they experienced some daytime sleepiness.

Due to dropouts and scheduling issues, there was some inconsistency in the counterbalancing of driving scenario used. Thirteen participants completed driving scenario A on their first study day (8 under normal sleep, 5 under sleep restriction) and 8 participants completed driving scenario B on their first study day (4 under normal sleep, 4 under sleep restriction).

### 5.3.2. Sleep Duration

Participants slept for an average of 453 minutes ( $SD = 60$ ) prior to the NS condition. Sleep was restricted to an average of 299 minutes ( $SD = 11$ ) during the SR condition.

### 5.3.3. Subjective Sleepiness

Participants were asked to rate their sleepiness using the KSS before, after and throughout each drive. Participants felt significantly sleepier following sleep restriction ( $M = 5.8$ ,  $SEM = 0.3$ ) compared with after a normal night of sleep ( $M = 3.7$ ,  $SEM = 0.3$ ). After the drive participants reported significantly greater effort to stay awake and experienced stronger sleep related eye symptoms (see Table 5-2).

Table 5-2

*Post drive measures of subjective sleepiness where higher scores indicate greater effort to stay awake*

Measure	NS <i>M (SEM)</i>	SR <i>M (SEM)</i>	<i>t</i> statistic	df.	Significance <i>p (r)</i>
Heavy eyelids	2.00 (0.21)	3.33 (0.27)	5.10	18	< .001 (.77)***
Difficulty keeping eyes open	1.48 (0.16)	2.57 (0.24)	4.11	20	.001 (.68)**
Difficulty focusing	1.86 (0.24)	2.95 (0.23)	3.75	20	.001 (.64)**
Eye strain	2.00 (0.23)	2.95 (0.29)	2.50	19	.022 (.50)*
Effort to stay awake	2.21 (0.29)	4.32 (0.34)	5.21	18	< .001 (.78)***
KSS (mean per drive)	3.73 (0.27)	5.75 (0.27)	7.18	20	< .001 (.85)***

*Note.* NS = Normal Sleep; SR = Sleep Restriction; M = mean; SEM = standard error of mean; df = degrees of freedom. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Participant scores on the NASA-TLX indicated they did not perceive the normal sleep drive ( $M = 199$ ,  $SE = 18$ ) as significantly less demanding than the sleep restriction drive ( $M = 241$ ,  $SE = 19$ ),  $t(20) = 1.76$ ,  $p = .09$ .

### 5.3.4. Change Detection Accuracy

Detection of changes following blackouts are presented as mean percentage correct within each environment and safety relevance level. Missing data occurred for 5.8% of events. Data was lost due to simulator error (failure to correctly display the event) and on some occasions where the change trial was presented but the participant did not respond. Six non-responses occurred under the sleep restriction condition (from 4 participants), 3 non-responses occurred under normal sleep from 3 participants. Table 5-3 displays the means and standard deviations of the high and low safety changes per driving environment expressed as percentages.

Table 5-3

*Correct responses to blackout events by driving environment, safety relevance and sleep condition*

Event Type	Normal Sleep <i>M (SD)</i>	Sleep Restriction <i>M (SD)</i>
Urban change-present: low safety relevance	28% (25)	25% (26)
Urban change-present: high safety relevance	60% (19)	54% (23)
Rural change-present: low safety relevance	52% (19)	47% (29)
Rural change-present: high safety relevance	87% (18)	94% (13)
Urban change-absent	88% (16)	93% (12)
Rural change-absent	93% (14)	92% (17)

Factorial repeated-measures ANOVA of change trials revealed that safety relevance and driving environment both had a main effect on change detection. Participants correctly detected more changes in the rural environment than in the urban environment,  $F(1,20) = 94.92$ ,  $p < .001$ . Participants also identified more high safety relevance changes than low relevance changes,  $F(1,20) = 66.63$ ,  $p < .001$ . Sleep restriction did not impact participants' ability to detect changes,  $F(1,20) = 0.20$ ,  $p = .661$ . A Bonferroni adjustment was applied to pairwise comparisons.

There was also a significant interaction between the driving environment and the safety relevance of the change,  $F(1,20) = 4.67$ ,  $p = .043$ . The improvement in detecting changes in rural compared with urban environments was more pronounced for high safety relevance changes, while driving environment made less of an impact on the number of low safety changes identified. This would be consistent with drivers becoming more vigilant in general when in urban environments but particularly more vigilant of changes that have greater potential to impact their safety.

### **5.3.5. Unexpected Change Detection Accuracy**

Table 5-4 indicates the percentage of correctly identified 'unexpected' changes during each condition. Change to the language of the signs was most noticed, followed by the lead vehicle changing colour and the number of pedestrians crossing behind the bus increasing. No participants under either condition noticed the change to road hatching colour. McNemar's test of consistency in responses indicated that there were no significant differences in the number of correctly identified 'unexpected' changes between the NS and SR conditions.

Table 5-4

*Percentage of participants correctly identifying unexpected changes levels by sleep condition.*

Unexpected Event	NS	SR	McNemar's test significance
Road hatching change colour	0%	0%	n/a
Lead vehicle change colour	33%	24%	$p = .500$ (1-tailed) $n = 20$
Text signs change to German	76%	67%	$p = .500$ (1-tailed) $n = 20$
One pedestrian changes to five	14%	5%	$p = .250$ (1-tailed) $n = 18$

*Note.* NS = Normal Sleep; SR = Sleep Restriction.

## 5.4. Discussion

The aim of the current study was to examine the impact of sleep loss on drivers' change detection ability for both expected changes (i.e., following screen blackout) and unexpected changes (i.e., occurring during a naturalistic occlusion of the scene) while driving in urban and rural environments. All participants were experienced, drivers who drove at least weekly. The results revealed that sleep loss did not significantly impact accuracy when detecting changes following screen blackouts. There is some suggestion that sleep loss reduces ability to identify unexpected changes; however, the difference was not statistically significant due to the small sample and effect size. Safety relevance and environment are more influential to change detection ability than sleep loss.

### 5.4.1. Effect of Sleep Loss

Despite feeling sleepier, requiring greater effort to stay awake, and experiencing an increase in sleep-related eye symptoms following a night of reduced sleep (5 hours), participants in the current study were not significantly impaired in change detection accuracy, relative to how they performed after a normal night of sleep (8 hours). This finding is consistent with Experiment 2, which also found no effect of sleep loss on change detection accuracy (but did find some more subtle effects in terms of both RT and ocular behaviour). However, the results of Experiments 2 and 3 are seemingly inconsistent with previous research using other tasks, which has consistently shown that sleep loss impairs attention and vigilance (Belenky et al., 2003; Dinges et al., 1997; Van Dongen et al., 2003).

One explanation for this discrepancy is that there are fundamental differences in both the task demands and the way that performance is measured and assessed. To our knowledge no other studies have specifically considered the impact of sleep loss on change blindness. Previous research examining the effects of sleep loss on attention has predominantly operationalised "attention" by measuring reaction time to simple, monotonous stimuli. A common task is the psychomotor vigilance task (PVT), which requires participants to monitor a screen and press a button as soon as a simple visual stimulus (e.g., a light) appears (Dinges et al., 1997). For the majority of trials reaction time is maintained at levels similar to performance when alert, however, intermittent trials are impaired. Sleep-related impairments in PVT usually manifest as an increase in lapses (i.e., failure to respond when the light appears) and increased variability in reaction time; for instance, the slowest 10% of reaction times will be significantly slower

following sleep loss, even if the median time is unchanged. As such, the impairment related to sleep-loss may be described as instability in performance, rather than gradual decline. This may explain why RT was impaired for some images in Experiment 2 but accuracy was not. Because the flicker sequence in Experiment 2 continued alternating for up to 30 s, the trial simply continued even if the participant became inattentive, so brief lapses of attention would manifest as an increase in RT rather than a decrease in accuracy. Importantly, however, Experiment 2 provided substantially more opportunities to measure attentional lapses, as each session contained 40 change-present trials, and participants' attentional performance was effectively monitored continuously throughout the 20-minute session. In contrast, Experiment 3 included only 16 events (12 following blackouts, 4 unexpected changes) in which aspects of the visual environment changed. With fewer data points it is less likely that a lapse would be recorded. To truly capture this effect it is necessary to record many trials which was not possible within the scope of Experiment 3.

For three of the unexpected changes a greater proportion of participants noticed the change while alert than when sleepy. Although the differences were not statistically significant, this could mean that sleep loss has greater influence on ability to detect changes when participants are not cued to respond. That is, the brief blackout period served as a cue to participants that something may have just changed, which would prompt them to make a special effort to be attentive immediately following the blackout, whereas detecting unexpected and uncued changes required participants to remain attentive throughout the entire drive. Furthermore, previous research has demonstrated that repeatedly driving the same route impairs drivers' ability to recognise some changes such as to road signs or road side buildings (Charlton & Starkey, 2013; Harms & Brookhuis, 2016; Martens & Fox, 2007). Sleep-related crashes most commonly occur within 5 km of the intended destination (Armstrong et al., 2013), and are therefore likely to be on familiar roads. The current work deliberately induced some familiarity by requiring participants to drive several laps of a looped route. However, future research may wish to consider whether sleepiness interacts with familiarity to increase the potential for change blindness by using a longer task that exposes participants repeatedly to the same segments of a road network.

#### ***5.4.2. Effect of Safety Relevance***

One of the main differences that emerged from the analyses was the impact of safety relevance. Specifically, high safety relevance changes were more accurately detected than low safety relevance changes. This finding is in line with those from Experiment 1. Previous research has demonstrated that changes to task relevant targets are detected more efficiently than changes to task irrelevant targets (Galpin et al., 2009; Lee et al., 2007; Mueller & Trick, 2013; Shinoda et al., 2001; Velichkovsky et al., 2002). In the context of simulated driving, safety relevance is one aspect of task relevance, as changes with higher safety relevance have greater immediate relevance to the driving task.

Whereas in Experiment 1 safety relevance mainly had an effect on change detection in urban scenes, in Experiment 3 safety relevance had a significant impact on change detection during both urban and rural driving. In Experiment 1 it was postulated that rural scenes were easier to process, which meant that participants had additional capacity to process change targets with lower safety relevance, thereby reducing the impact of safety relevance (in contrast with urban scenes, where the impact of safety relevance was much larger).



One notable difference between Experiment 1 and 3 is that during Experiment 3 participants were moving through the environment rather than observing a static image. The high speed of travel on rural roads could have increased the task difficulty; thus the increased processing required to process changes in the environment while moving at high speed may have resulted in a greater impact of safety relevance to change detection ability.

#### ***5.4.3. Effect of Driving Environment***

As with Experiment 1 and 2, participants were more accurate at detecting changes in rural than urban environments. These differences are most likely attributable to the fact that urban scenes involve greater visual clutter and complexity. To our knowledge, no previous research has compared change detection in urban and rural scenes in the same way but the replication of findings between each experiment in this series suggests this is a robust outcome.

#### ***5.4.4. Simulator Methodology Limitations***

The findings of the current work are limited by the small number of change-detection trials completed. Compared with Experiment 1 and 2, which used static images, in Experiment 3 it was necessary to reduce the number of change events that occurred as each event had to be programmed and scripted within the simulator software. In order to occlude cued changes, a new programming protocol had to be developed under which changes could occur out of participants view. This was very demanding in terms of the programming required and the resulting scenarios imposed unusually high processing demands on the simulator software and hardware.

The experimental protocol was designed to measure both accuracy and RT of responses following blackouts. RT was recorded using a custom-made in-vehicle touch screen device which participants could press as soon as they identified a change. Similar systems have previously been successfully used to record participants' RT to identify target vehicles while driving on straight roads with minimal traffic and relatively low speeds (e.g., Beanland et al., 2014a). However, in the current study the driving task was more demanding, so in practice participants' first priority was maintaining safe control of the vehicle. During data collection it became apparent that button presses to high safety relevance changes in particular were not necessarily representative of the point when the participant detected the change. For example, for the changes involving an oncoming vehicle appearing in the participant's lane, a participant might immediately detect this event and respond by braking in order to reduce the potential for a collision, and therefore delay responding via the button press.

For cases where the button press RTs may not represent the true RT, simulator recorded measures of vehicle control (e.g., speed, road positioning) could provide a proxy measure of detection and response. However, the SCANer™ studio software failed to record these metrics for several participants and because of the volume of missing data, the limited data that was available for the participants with intact data sets was not analysed. It is possible that the recording failures were due to the high processing demands placed on the software by the change detection paradigm. Similarly, these computer processing demands meant it was not possible to integrate the eye-tracking system with the simulator scenario, so eye movements could not be reliably tracked during the simulator scenarios. Overall, a driving simulator study of this nature is highly resource intensive, both in terms of the

programming and the technical resources required, and future research in this area would require substantially more resources, over a longer timeframe, in order to more comprehensively measure change detection together with vehicle measures and eye movements. Given the results from the current experimental series, it is likely that the addition of vehicle measures and eye-movements will offer valuable additional insights into drivers' change detection.

#### ***5.4.5. Summary***

Overall the results of Experiment 3 indicate that change detection accuracy is not affected by sleep loss. Both the driving environment in which the change occurs and the safety relevance of the change are better predictors of whether a change will be accurately and rapidly identified. Specifically, drivers are more efficient at detecting changes in rural environments and those of greater safety relevance.

Notably, the impact of safety relevance was greater for changes in rural than urban environments. This may be because the higher travel speed through rural environments reduced processing time therefore making the low safety relevant changes less noticeable.

Finally, accuracy for identifying unexpected changes is much lower than for changes when participants are cued to respond. This finding is consistent with basic laboratory experiments comparing detection of expected and unexpected changes (Beck et al., 2007). In real-world driving some hazards develop in a way that may provide drivers with cues to direct their attention (or remind them to search for hazards). This type of developing hazard is often researched within road safety particularly in relation to differences in anticipation and situation awareness between experienced and novice drivers (Underwood, 2007). However, other hazards may change subtly and be unexpected. Familiarity appears to exacerbate change blindness in this situation for particular objects such as signs and buildings (Charlton & Starkey, 2013) and current results provide some suggestion that sleepiness may have greater influence on this type of change detection. Detection and response to unexpected changes is currently under researched. Future work may wish to consider what factors improve drivers' ability to detect these changes.

## 6. Summary and Conclusions

The current research program comprised a series of three experiments designed to examine factors that influence drivers' efficiency of change detection, with a particular emphasis on how sleep loss impacts change detection.

Overall, the research findings provide only limited evidence that change detection while driving is impaired by sleep loss. Specifically, in both Experiments 2 and 3 accuracy of detecting changes was not significantly affected by sleepiness; however, some more sensitive performance measures showed a small degree of impairment. In Experiment 2, participants required longer to detect changes in urban scenes when they were sleepy, compared with when they were alert, and in Experiment 3 there was a non-significant reduction in detection of unexpected changes. The impact of even small increases in RT could potentially be catastrophic under certain circumstances; the 370 ms increase in RT would equate to an extra 6.2 m at a travel speed of 60 km/h. In a critical situation, this distance could represent the difference between a collision and a near miss.

The current work explored topics that had not been investigated previously, specifically by examining how sleep restriction affects complex visual attention tasks. It appears that the impact of sleep loss on change detection is subtle, with most analyses not indicating any statistically significant differences. However, as with any "null finding", the lack of significance should be interpreted with caution. Specifically, the lack of a statistically significant difference does not mean that two conditions are the same. Both Experiments 2 and 3 used relatively small samples of participants, which is a practical limitation imposed by the resources available. The repeated-measures nature of sleep research also often results in attrition of participants, as some individuals decide after one or two sessions that they prefer not to continue, and in Experiment 3 the dropout issue was further exacerbated as several participants experienced simulator sickness and were forced to withdraw.

Another issue that affected Experiment 3 disproportionately was the number of measurements, as there were relatively few trials of the change detection task. Previous research has consistently demonstrated that sleep loss impairs attention (e.g., Belenky et al., 2003; Dinges et al., 1997; Van Dongen et al., 2003); however, the types of tasks that show robust effects of sleep loss have many different characteristics when compared with change detection tasks. The most common test paradigm for investigating sleep loss and attention is the PVT. Impairment at this simple reaction time task usually manifest as an increase in the number of lapses (i.e., failure to respond when the light appears) and increased variability in reaction time; for instance, the slowest 10% of reaction times will be significantly slower following sleep loss, even if the median time is unchanged. It is possible that the current change-blindness paradigm is not sufficiently sensitive to be able to identify occasional lapses. This is particularly likely to be true for Experiment 3, which had relatively few change detection events. Future research should therefore aim to include a much larger number of trials (which may, in practice, preclude the possibility of conducting the study in a driving simulator) to ascertain whether the relatively small number of measurements is contributing to the null effect. By increasing the number of measurements, it may also be possible to conduct statistical analyses that compare measures other than the mean

performance, for instance, looking at the fastest and slowest 10% of responses in a manner similar to typical PVT analyses.

Aside from differences in the number and type of measurements used in PVT studies compared with change detection studies, there are also fundamental differences in the nature of the tasks that could contribute to the non-significant effects of sleep loss observed in Experiments 2 and 3. Specifically, the PVT is a very simple response task, in which participants respond as quickly as possible to the presence of a light that appears intermittently. Change detection tasks are much more complex, involving not only visual attention but also working memory, and the stimuli used were richer and more interesting. In general, the tasks used in Experiments 2 and 3 are less monotonous and more engaging than both the PVT and the types of driving simulator scenarios that usually reveal robust effects of sleep loss. For example, in Experiment 2 the flicker task changed the screen presentation every 500 ms, alternating between the photographic image and a blank screen. It may be that the blank screen helped participants focus their attention, as they could “zone out” briefly during blanks and refocus when the image reappeared. Similarly, for Experiment 3 the driving scenario included complex urban roads in addition to the more traditionally studied monotonous rural roads, and participants encountered vehicular traffic, pedestrians and intersections more often than in a typical driver sleepiness study. It is possible that the urban driving environment was sufficiently alerting to preserve change-detection performance despite the pressure of sleep loss.

Driver sleepiness contributes to an estimated 15-30% of all crashes (Åkerstedt, 2000; Connor, 2009; Horne & Reyner, 1995). Despite recent research indicating that driver sleepiness occurs frequently in urban environments (Armstrong et al., 2011, 2013; Filtiness et al., under review), most experimental research has focused on how sleepiness affects driving performance in monotonous rural environments. It is possible that the relationship between driver sleepiness and crash risk may differ between urban and rural environments, due to differences in travel speeds and road characteristics. Rural roads are more likely to be monotonous, high-speed roads and sleepy drivers may experience attentional lapses because the environment is insufficiently stimulating. Then, because their travel speed is high, the consequences of this attentional lapse may be more severe, resulting in an extreme out-of-lane event or even a crash. On lower-speed urban roads, sleepy drivers may be less likely to experience attentional lapses in the first place, as the environment itself is more stimulation, and if they do experience a lapse the consequences may be less severe due to the lower travel speed. Thus attentional lapses in urban areas may manifest more subtly, as near misses or increases in variability of the driver’s lateral positioning on the road. Interestingly, although the consequences of attentional lapses are likely to be less severe, the results of the current research suggest that attentional lapses are considerably more common in urban environments. Across all three experiments, participants demonstrated relatively more difficulty when required to detect changes in urban driving scenes, compared with similar changes in rural environments. Overall, in future research it would be worth further examining the impact of driver sleepiness in urban environments, but researchers should be mindful of the fundamental differences between urban and rural driving that may influence the pattern of results obtained.

Although the current research found that sleep loss had minimal impact on change detection in driving scenes, it identified several other factors that do profoundly influence drivers’ accuracy and speed of

change detection. Factors that consistently influenced change detection in all three experiments were the type of driving environment, type of object that changes, and the safety relevance of the change that occurred. Participants were consistently better at detecting changes with higher safety relevance, even though all changes involved objects that were “task relevant”. This finding is notable, as previous research has conceptualised task relevance as dichotomous and has compared broad categories of “relevant” and “irrelevant” changes (Galpin et al., 2009; Mueller & Trick, 2013; Velichkovsky et al., 2002; Zhao et al., 2014). The current research suggests that relevance is on a continuum, and that change detection performance may follow that continuum, so that changes become relatively more or less difficult to detect as they become less or more relevant to the task at hand.

In addition to the consistent effect of safety relevance on change detection, the type of object that changed appears to have a large effect on change detection. Obviously in the real world, there is a correlation object type and safety relevance, as some objects pose greater hazards while driving. As demonstrated in Experiment 1, the effect of change target type was significant even when controlling for safety relevance, suggesting that participants preferentially attended to certain aspects of the environment. In particular, accuracy, RT and eye-tracking results indicated that participants were relatively more attentive to objects that could plausibly move or change in the real world, including road users, animals and traffic lights, and paid relatively little attention to static objects such as signs and trees.

The fact that drivers are relatively inattentive to signs is alarming, given that governments commonly rely on signs to convey important road safety information, including appropriate travel speeds. Future research should carefully examine how road signs are designed and placed, with the aim of identifying conditions that enhance drivers’ attentiveness to signs. It may also be advisable to remove redundant signs and to restrict placement of irrelevant signs (i.e., not driving related) near the roadside, so that roadside signs become more relevant to drivers. Finally, as many jurisdictions use dynamic and variable message signs to convey particularly urgent and time-sensitive information, it would be interesting in future research to examine whether drivers are similarly inattentive to dynamic signs.

Finally, the driving environment in which changes occurred had a substantial impact on change detection performance. In all three experiments, drivers were better at detecting changes in rural environments compared with urban environments, even though the characteristics of the change (e.g., type of object that changed, safety relevance and nature of the change) were matched between environments. This effect is not surprising, as urban environments are more cluttered and therefore provide more visual distraction, so it is more difficult to identify specific objects of interest. However, it is notable that there was a slight discrepancy between the results obtained using photographic images (Experiment 1) compared with simulated driving (Experiment 3). Specifically, safety relevance had a much greater impact on change detection in rural areas in Experiment 3, whereas in Experiment 1 safety relevance had minimal impact. The most likely explanation is that visually scanning static photographs of rural scenes is considerably easier than conducting the same type of scanning while driving through the environment at 100 km/h. In other words, adding the requirement to navigate through the simulated environment increased the task demands, which resulted in impairments to change detection.

This discrepancy between Experiments 1 and 3 highlights the need for future research to confirm the extent to which experimental findings can be generalised from computer-based experiments to real world tasks. Our findings are consistent with previous change detection research by Velichkovsky et al. (2002), who also found a greater degree of change blindness in dynamic, simulated scenarios, compared to in static experiments. Driving simulators are an appropriate medium with which to conduct dynamic change blindness studies; however, there are many challenges to be overcome on how best to adapt change detection paradigms into a driving simulator, and substantial resources are necessary to overcome both programming and eye tracking difficulties.

In summary, the current research confirms that drivers often experience change blindness, which refers to difficulty in detecting when elements of their environment have been changed. Change blindness is most likely to occur in urban environments, when the change has little potential to influence driving safety, or when it involves a static or fixed object such as a road sign. Based on the current research, it appears that sleep loss does not have a large or universal effect on drivers' change detection ability, but rather that it may have small and subtle effects on some aspects of change detection, such as the time required to detect changes in urban environments (Experiment 2) or ability to detect unexpected changes (Experiment 3). Future research should focus on better understanding these effects, particularly focusing on the potential for driver sleepiness to adversely impact change detection in urban driving and/or in unexpected situations.

## 7. References

- ACT Government. (2011). *Road Safety Strategy 2011-20*. Canberra, Australia: ACT Government, Justice and Community Services.
- Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52(1-2), 29-37.
- Åkerstedt, T. (2000). Consensus statement: fatigue and accidents in transport operations. *Journal of Sleep Research*, 9(4), 395-395.
- Anderson, C., & Horne, J. A. (2013). Driving drowsy also worsens driver distraction. *Sleep Medicine*, 14(5), 466-468.
- Anund, A., Kecklund, G., Peters, B., Forsman, A., Lowden, A., & Åkerstedt, T. (2008). Driver impairment at night and its relation to physiological sleepiness. *Scandinavian Journal of Work, Environment & Health*, 34(2), 142-150.
- Armstrong, K., Filtress, A. J., Watling, C. N., Barraclough, P., & Haworth, N. (2013). Efficacy of proxy definitions for identification of fatigue/sleep-related crashes: An Australian evaluation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 21, 242-252.
- Armstrong, K., Haworth, N., Watling, C., & Trenorden, J. (2011). *Improving the identification of fatigue-related crashes in the Australian Capital Territory*. Report prepared for NRMA-ACT Road Safety Trust and the Roads and Traffic Authority, New South Wales. Brisbane, Australia: Centre for Accident Research & Road Safety–Queensland.
- Attewell, R., Lock, S., Dobbie, K., & Walker, M. (2001). Dead tired: Fatigue related crashes on National Route 39. In: *Proceedings of the Australasian Road Safety Research, Policing and Education Conference*, Melbourne, Victoria, Australia.
- Beanland, V., Fitzharris, M., Young, K. L., & Lenné, M. G. (2013). Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Accident Analysis & Prevention*, 54, 99-107.
- Beanland, V., Lenné, M. G., & Rößger, L. (2015). Psychological factors in seeing motorcycles. In L. Rößger, M. G. Lenné & G. Underwood (Eds.), *Increasing Motorcycle Conspicuity: Design and Assessment of Interventions to Enhance Rider Safety* (pp. 21-49). Aldershot, UK: Ashgate.
- Beanland, V., Lenné, M. G., & Underwood, G. (2014a). Safety in numbers: Target prevalence affects the detection of vehicles during simulated driving. *Attention, Perception, & Psychophysics*, 76(3), 805-813.
- Beanland, V., Sellbom, M., & Johnson, A. K. (2014b). Personality domains and traits that predict self-reported aberrant driving behaviours in a southeastern US university sample. *Accident Analysis & Prevention*, 72, 184-192.
- Beck, M. R., Levin, D. T., & Angelone, B. (2007). Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*, 16(1), 31-51.



- Belenky, G., Wessensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B., & Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research*, 12(1), 1-12.
- Bridger, R. S., Johnsen, S. Å. K., & Brasher, K. (2013). Psychometric properties of the Cognitive Failures Questionnaire. *Ergonomics*, 56(10), 1515-1524.
- 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.
- Caird, J. K., Edwards, C. J., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors*, 47(2), 235-249.
- Chapman, P. R., & Underwood, G. (1998). Visual search of dynamic scenes: Event types and the role of experience in viewing driving situations. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 369-392). Oxford, UK: Elsevier.
- Charlton, S. G., & Starkey, N. J. (2013). Driving on familiar roads: Automaticity and inattention blindness. *Transportation Research Part F: Traffic Psychology and Behaviour*, 19, 121-133.
- Charlton, S. G., Starkey, N. J., Perrone, J. A., & Isler, R. B. (2014). What's the risk? A comparison of actual and perceived driving risk. *Transportation Research Part F: Traffic Psychology and Behaviour*, 25, Part A, 50-64.
- Clark, B., & Warren, N. (1939). The effect of loss of sleep on visual tests. *American Journal of Optometry*, 16, 80-95.
- Connor, J. L. (2009). The role of driver sleepiness in car crashes: a review of the epidemiological evidence. In *Drugs, Driving and Traffic Safety* (pp. 187-205). Basel, Switzerland: Birkhäuser.
- Dawson, D., & Reid, K. (1997). Fatigue, alcohol and performance impairment. *Nature*, 388(6639), 235-235.
- De Gennaro, L., Ferrara, M., Curcio, G., & Bertini, M. (2001). Visual search performance across 40 h of continuous wakefulness: Measures of speed and accuracy and relation with oculomotor performance. *Physiology & Behavior*, 74(1), 197-204.
- De Gennaro, L., Ferrara, M., Urbani, L., & Bertini, M. (2000). Oculomotor impairment after 1 night of total sleep deprivation: a dissociation between measures of speed and accuracy. *Clinical Neurophysiology*, 111(10), 1771-1778.
- Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., Aptowicz, C., & Pack, A. I. (1997). Cumulative sleepiness, mood disturbance and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*, 20(4), 267-277.
- Edwards, C. J., Caird, J. K., & Chisholm, S. L. (2008). The effect of driving experience on change blindness at intersections: Decision accuracy and eye movement results. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(18), 1243-1247.

- Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 97-120.
- Falkmer, T., & Gregersen, N. P. (2005). A comparison of eye movement behavior of inexperienced and experienced drivers in real traffic environments. *Optometry & Vision Science*, 82(8), 732-739.
- Feil, A., & Mestre, J. P. (2010). Change blindness as a means of studying expertise in physics. *Journal of the Learning Sciences*, 19(4), 480-505.
- Filtness, A. J., Reyner, L. A., & Horne, J. A. (2012). Driver sleepiness—Comparisons between young and older men during a monotonous afternoon simulated drive. *Biological Psychology*, 89(3), 580-583.
- Filtness, A. J., Armstrong, K. A., Watson, A., & Smith, S. A. (under review). Sleep-related vehicle crashes on low speed roads. Manuscript submitted for publication.
- Fors, C., Ahlström, C., & Anund, A. (2013). *Simulator validation with respect to driver sleepiness and subjective experiences: final report of the project SleepEYE II, part 1*. Retrieved from <http://www.diva-portal.org/smash/get/diva2:674216/FULLTEXT01.pdf>
- Forsman, P. M., Vila, B. J., Short, R. A., Mott, C. G., & Van Dongen, H. P. (2013). Efficient driver drowsiness detection at moderate levels of drowsiness. *Accident Analysis & Prevention*, 50, 341-350.
- Forster, S., & Lavie, N. (2007). High perceptual load makes everybody equal eliminating individual differences in distractibility with load. *Psychological Science*, 18(5), 377-381.
- Fransson, P. A., Patel, M., Magnusson, M., Berg, S., Almbladh, P., & Gomez, S. (2008). Effects of 24-hour and 36-hour sleep deprivation on smooth pursuit and saccadic eye movements. *Journal of Vestibular Research*, 18(4), 209-222.
- Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(2), 179-185.
- Goldich, Y., Barkana, Y., Pras, E., Zadok, D., Hartstein, M., & Morad, Y. (2010). The effects of sleep deprivation on oculomotor responses. *Current Eye Research*, 35(12), 1135-1141.
- Haavisto, M.-L., Porkka-Heiskanen, T., Hublin, C., Härmä, M., Mutanen, P., Müller, K., Virkkala, J., & Sallinen, M. (2010). Sleep restriction for the duration of a work week impairs multitasking performance. *Journal of Sleep Research*, 19(3), 444-454.
- Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J.G., Wahde, M., & Åkerstedt, T., (2013). Sleepy driving on the real road and in the simulator—a comparison. *Accident Analysis & Prevention* 50, 44-50.
- Harms, I. M., & Brookhuis, K. A. (2016). Dynamic traffic management on a familiar road: Failing to detect changes in variable speed limits. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 37-46.

- Haque, M. M., & Washington, S. (2013). Effects of mobile phone distraction on drivers' reaction times. *Journal of the Australasian College of Road Safety*, 24(3), 20-29.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183.
- Herman, J., Kafoa, B., Wainiqolo, I., Robinson, E., Mccaig, E., Connor, J., Jackson, R., & Ameratunga, S. (2014). Driver sleepiness and risk of motor vehicle crash injuries: A population-based case control study in Fiji (trip 12). *Injury*, 45(3), 586-591.
- Horne, J. (2012). Working throughout the night: Beyond 'sleepiness'—impairments to critical decision making. *Neuroscience & Biobehavioral Reviews*, 36(10), 2226-2231.
- Horne, J. A. (1975). Binocular convergence in man during total sleep deprivation. *Biological Psychology*, 3(4), 309-319.
- Horne, J. A., & Reyner, L. A. (1995). Sleep related vehicle accidents. *BMJ*, 6979, 565-567.
- Horne, J. A., & Reyner, L. A. (1996). Counteracting driver sleepiness: Effects of napping, caffeine, and placebo. *Psychophysiology*, 33(3), 306-309.
- Jackson, M. L., Croft, R. J., Kennedy, G. A., Owens, K., & Howard, M. E. (2013). Cognitive components of simulated driving performance: Sleep loss effects and predictors. *Accident Analysis & Prevention*, 50, 438-444.
- Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattention blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5), 529-546.
- Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep*, 14(6), 540-545.
- Josephs, E., Drew, T., & Wolfe, J. (2016). Shuffling your way out of change blindness. *Psychonomic bulletin & review*, 23(1), 193-200.
- Koustanai, A., Van Elslande, P., & Bastien, C. (2012). Use of change blindness to measure different abilities to detect relevant changes in natural driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(3), 233-242.
- Lajunen, T., Parker, D., & Summala, H. (2004). The Manchester Driver Behaviour Questionnaire: A cross-cultural study. *Accident Analysis & Prevention*, 36(2), 231-238.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3), 451-468.
- Lawton, R., Parker, D., Manstead, A. S. R., & Stradling, S. G. (1997). The role of affect in predicting social behaviors: The case of road traffic violations. *Journal of Applied Social Psychology*, 27(14), 1258-1276.
- Lee, Y.-C., Lee, J. D., & Ng Boyle, L. (2007). Visual attention in driving: The effects of cognitive load and visual disruption. *Human Factors*, 49(4), 721-733.

- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13-22.
- Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Predicting driver drowsiness using vehicle measures: Recent insights and future challenges. *Journal of Safety Research*, 40(4), 239-245.
- Ma, Y., Mazumdar, M., & Memtsoudis, S. G. (2012). Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Regional Anesthesia and Pain Medicine*, 37(1), 99-105.
- Marchetti, L. M., Biello, S. M., Broomfield, N. M., Macmahon, K. M. A., & Espie, C. A. (2006). Who is pre-occupied with sleep? A comparison of attention bias in people with psychophysiological insomnia, delayed sleep phase syndrome and good sleepers using the induced change blindness paradigm. *Journal of Sleep Research*, 15(2), 212-221.
- Martens, M. H., & Fox, M. R. J. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(6), 476-492.
- Matthews, R. W., Ferguson, S. A., Zhou, X., Kosmadopoulos, A., Kennaway, D. J., & Roach, G. D. (2012). Simulated driving under the influence of extended wake, time of day and sleep restriction. *Accident Analysis & Prevention*, 45, 55-61.
- Mattsson, M. (2012). Investigating the factorial invariance of the 28-item DBQ across genders and age groups: An exploratory structural equation modeling study. *Accident Analysis & Prevention*, 48, 379-396.
- McCarley, J. S., Vais, M. J., Pringle, H., Kramer, A. F., Irwin, D. E., & Strayer, D. L. (2004). Conversation disrupts change detection in complex traffic scenes. *Human Factors*, 46(3), 424-436.
- McClelland, L. E., Pilcher, J. J., & Moore, D. D. (2010). Oculomotor measures as predictors of performance during sleep deprivation. *Aviation, Space, and Environmental Medicine*, 81(9), 833-842.
- McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, 22, 563-581.
- Mueller, A. S., & Trick, L. M. (2013). *Effect of driving experience on change detection based on target relevance and size*. Paper presented at the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44, 369-378.
- Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics*, 38(5), 1036-1048.
- Philip, P., Sagaspe, P., Moore, N., Taillard, J., Charles, A., Guilleminault, C., Bioulac, B., (2005). Fatigue, sleep restriction and driving performance. *Accident Analysis & Prevention* 37 (3), 473-478.

- Philip, P., Sagaspe, P., Lagarde, E., Leger, D., Ohayon, M.M., Bioulac, B., Boussuge, J., & Taillard, J., (2010). Sleep disorders and accidental risk in a large group of regular registered highway drivers. *Sleep Medicine*, 11(10), 973-979.
- Radun, I., Ohisalo, J., Radun, J., Wahde, M., & Kecklund, G. (2013). Driver fatigue and the law from the perspective of police officers and prosecutors. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 159-167.
- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, 12(1), 48-55.
- Rensink, R. A., O'Regan, K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368-373.
- Richard, C. M., Wright, R. D., Ee, C., Prime, S. L., Shimizu, Y., & Vavrik, J. (2002). Effect of a concurrent auditory task on visual search performance in a driving-related image-flicker task. *Human Factors*, 44(1), 108-119.
- Rowland, L. M., Thomas, M. L., Thorne, D. R., Sing, H. C., Krichmar, J. L., Davis, H. Q., Balwinski, S. M., Peters, R. D., Kloeppel-Wagner, E., Redmond, D. P., Alicandri, E., & Belenky, G. (2005). Oculomotor responses during partial and total sleep deprivation. *Aviation, Space, and Environmental Medicine*, 76(7), C104-C113.
- Russo, M., Thomas, M., Thorne, D., Sing, H., Redmond, D., Rowland, L., Johnson, D., Hall, S., Krichmar, J., & Balkin, T. (2003). Oculomotor impairment during chronic partial sleep deprivation. *Clinical Neurophysiology*, 114(4), 723-736.
- Russo, M., Thomas, M., Sing, H., Thorne, D., Balkin, T., Wesensten, N., Redmond, D., Welsh, A., Rowland, L., & Johnson, D. (1999). Saccadic velocity and pupil constriction latency changes in partial sleep deprivation, and correlations with simulated motor vehicle crashes. *Sleep*, 22(Suppl 1), S297-298.
- Sagberg, F., Road accidents caused by drivers falling asleep. (1999). *Accident Analysis & Prevention*, 31(6), 639-649.
- Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research*, 41(25-26), 3535-3545.
- Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5, 644-649.
- Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, 50(8), 1235-1249.
- Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving: skill and awareness during inspection of the scene. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(2), 87-97.

- Van Dongen, H. P., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: dose-response effects on neurobehavioural functions and sleep physiology from chronic sleep restriction and total deprivation. *Sleep*, 26(2), 117-126.
- Van Dongen, H. P. A., Baynard, M. D., Maislin, G., & Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep*, 27(3), 423-433.
- Velichkovsky, B. M., Dornhoefer, S. M., Kopf, M., Helmert, J., & Joos, M. (2002). Change detection and occlusion modes in road-traffic scenarios. *Transportation Research Part F: Traffic Psychology & Behaviour*, 5(2), 99-109.
- Wallace, C. J. (2004). Confirmatory factor analysis of the cognitive failures questionnaire: evidence for dimensionality and construct validity. *Personality and Individual Differences*, 37(2), 307-324.
- Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, 7(1-3), 163-173.
- 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. *Accident Analysis & Prevention*, 42(4), 1232-1239.
- White, C. B., & Caird, J. K. (2010). The blind date: The effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident Analysis & Prevention*, 42(6), 1822-1830.
- Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15(4), 160-168.
- Zhao, N., Chen, W., Xuan, Y., Mehler, B., Reimer, B., & Fu, X. (2014). Drivers' and non-drivers' performance in a change detection task with static driving scenes: is there a benefit of experience? *Ergonomics*, 57(7), 998-1007.