The Australian 400-car Naturalistic Driving Study: Innovation in road safety research and policy

Regan, M.A[a]., Williamson, A[a]., Grzebieta, R[a]., Charlton[b], J., Lenne[b], M., Watson[c], B., Haworth, N[c]., Rakotonirainy, A[c]., Woolley, J[d]., Anderson, R[d]., Senserrick, T[a]., & Young, K[b]

Transport and Road Safety (TARS) Research, University of NSW, Sydney, Australia [a]
Monash University Accident Research Centre, Melbourne, Australia [b]
Centre for Accident Research and Road Safety, Queensland University of Technology, Brisbane, Australia [c]
Centre for Automotive Safety Research, University of Adelaide, Adelaide, Australia [d]

Abstract

This paper describes the design of Australia's first large-scale Naturalistic Driving Study (NDS), the aim of which is to understand what people *actually* do when they drive their cars in normal, impaired and safety-critical situations. In early 2014, each of 400 volunteer participants - 200 in NSW (Sydney and Wagga Wagga) and 200 in Victoria (Melbourne and Bendigo) - will have their own vehicle instrumented for 6 months with a Data Acquisition System (DAS) which will record continuously their driving behaviour (e.g. where they are looking), the behaviour of their vehicle (e.g. speed, lane position, headway), the behaviour of other road users with whom they interact (e.g. other drivers, motorcyclists, cyclists and pedestrians), and their interactions with the road infrastructure. Prior to participating, each driver will undergo an extensive testing regime that will enable the research team to correlate their personal data with on-road data. Drivers will a have valid Australian licence (provisional or full licence) and be aged between 17 and 70 years. Eight overarching research questions will guide the exploration of the data: (1) What collision risks are drivers exposed to?; (2) What collision risks do drivers expose themselves and others to?; (3) By how much do these exposures increase collision risk to them and to others?; (4) What are the factors that protect drivers against crash and injury?; (5) How do drivers modify their behaviour to adapt to conditions of increased and potential collision risk?; (6) How do people drive normally to minimise collision risk and avoid crashes?; (7) What are the factors that cause and contribute to crashes and safety-critical events?; and (8) What is the relationship between driver personal characteristics, driving behaviour and crash risk? The main outcomes of the study will be the formulation, based on the answers to these research questions, of new and improved evidence-based countermeasures for improving intersection safety, reducing speed-related crashes, reducing fatal and serious injury crashes involving vulnerable road users (pedestrians, cyclists and motorcyclists), reducing crashes involving fatigue and inattention, and improving the design and performance of collision warning and other intelligent vehicle safety technologies.

Introduction

In Australia, around 1400 people die and 33,000 are seriously injured on average each year on our roads (ATC, 2011). The estimated annual cost of road trauma to the Australian community is $AUD 27 billion (BITRE, 2009). While Australia's annual road fatality rate declined from 22.3 deaths per 100,000 people in 1980 to 6.1 in 2010 (ATC, 2011), the rate of decline has slowed markedly over the last decade. In 2009, Australia slipped to 16th out of 27 OECD countries in terms of road fatalities per capita (ATC, 2011), after being in the top 10 during the previous decade (ATSB 2002; 2005). The best performing countries (e.g., Sweden, the Netherlands) are presently achieving rates consistently below 4.0 fatalities per 100,000 people.

Australia's past success in road safety has been due, in large part, to the development of road safety strategies with prioritised interventions with a very strong evidence base. To date, this evidence base has been derived primarily from crash data collected by police, in-depth crash investigations, Coroners' and hospital data and from data from surveys on driver exposures to risk. However, these data sources are limited in the depth and quality of information they provide about driver and road user behaviour, which are major contributing factors in most collisions (Antin et al., 2011). Such data can often only be inferred, if at all, from available evidence after a crash or from surveys with confounding unknown self-reported biases (Gordon & Regan, 2012). Existing data collection methods in road safety in Australia rely on the limited post-crash accuracy and biases of driver and witness recall of events and on retrospective physical evidence from crash scenes - with little or no pre-crash information about other vehicles and road users involved.

The Naturalistic Driving Study (NDS) is a relatively new research method that has the potential to overcome many of these limitations. In an NDS, volunteer participants drive an instrumented vehicle (usually their own) for 6 to 12 months, or more, fitted with an unobtrusive Data Acquisition System (DAS) which continuously records their driving behaviour (e.g. where they are looking), the behaviour of their vehicle (e.g. speed, lane position) and the behaviour of other road users with whom they interact (e.g. other drivers, motorcyclists, cyclists and pedestrians) - in normal and safety-critical situations. Each DAS, depending on its capabilities, incorporates multiple sensors (video cameras, GPS, radar, accelerometers, etc.) – to provide a complete, second-by-second, picture of driver, vehicle and road user behaviour in all driving situations (Klauer et al., 2011). Awareness of being filmed has potential in such studies to constrain the expression of natural driving behaviours; however, NDS experience in the US demonstrates that drivers revert quickly back to their natural driving behaviours, within an hour or so (Dingus et al., 2006).

Advantages of the NDS

The NDS offers a new approach for understanding driver and vehicle behaviour in normal, impaired and safety-critical situations. Regan, Williamson, Grzebieta & Tao (2012, p 3) have highlighted some of the key advantages to be gained in running a large-scale NDS in Australia:

- **Exposure**: "New and more detailed data can be collected on driver exposure to a wider range of driver, vehicle, road, traffic and environmental factors that increase crash risk: for example, speeding; inappropriate vehicle positioning (e.g., tailgating); being distracted/inattentive; carrying passengers disobeying road signs and signals; driving at night; driving in poor weather; and driving in the country; etc.
- **Crash risk**. Having better exposure data makes it is possible to calculate odds ratios (relative risk) and population attributable risk percentages (proportion of crashes) for a much broader range of risky activities to which drivers expose themselves; and the derivation of odds ratios that are more situational-specific e.g. "the risk of traveling 5 km/h over the speed limit for a young driver driving at night on a country road, carrying a passenger.'
- **Near-Crashes.** Data on near-crashes, of which thousands occur regularly on Australian roads, but which are never reported, would be retrieved, yielding previously unavailable data on how near-crashes come about – e.g., between drivers and other road users (pedestrians, motorcycle riders, cyclists, car drivers, truck drivers, etc.) – and on how road users prevent near-crashes with other road users from becoming crashes. Such data can be used to, among other things, optimise training and education programs, and optimise the design of advanced driver assistance systems, the traffic management system, and the road environment.
- **Crashes**. Typically, there have been few crashes in naturalistic driving studies, because the number of vehicles/drivers involved in them has been in the order of hundreds (e.g., Klauer et al., 2006) rather than thousands (e.g., Antin et al., 2011). Nevertheless, the crashes that do occur reveal the truth – the factors that actually contribute to crashes. Furthermore, it is possible to discover what factors turn a near-crash into a crash, and hence to understand the critical differences between near-crashes and crashes. Such data can be used, for example, to optimise

the design of advanced driver assistance systems, the traffic management system, and the road environment.

- **Normative data**. As noted, the NDS provides normative data for international benchmarking that provides a confirmatory check on the applicability to Australia of other NDS data. The NDS makes it possible to gather fundamental data on how people drive – how they avoid crashes, navigate, maintain speed, adhere to traffic laws; stay within their lane; control the vehicle etc. – and how this varies according to age, experience, driver state (e.g. drunk, distracted), driver condition (demented, etc.) and other factors. These data can be used to build and refine models of driving behaviour, refine traffic micro-simulations, develop training and education programs, etc.
- **Violations**. The NDS can provide the road safety community with a better understanding of which traffic laws are violated, by whom (drivers and other road users), where, when (e.g., at night) and in what situations (e.g., when distracted, drunk). This data can be used to improve traffic laws, optimise Police enforcement regimes, and re-design traffic management practices, among other things.
- **Validation**. Objective data yielded by the NDS can be used to validate findings that emerge from self-report studies of driver behaviour, from on-site observational studies and simulators driven by humans. Drivers do not always do what they say they do (e.g., Myers et al., 2011) and not all findings from simulators are replicated in real traffic, and vice-versa (e.g., Olson et al, 2009). In the case of crashes, causal factors derived from the NDS can be compared with those cited in Police report forms to validate Police reports. Evidence from NDS studies in the US indicates that crashes recorded both by DAS units and Police are miscoded routinely by Police.
- **Evaluation**. Traditionally, Australia has placed great emphasis on the need to evaluate countermeasures, to determine whether they are effective and how they can be optimised. The NDS, suitably designed, can be used to derive new data that can be used to evaluate the effectiveness of new and existing countermeasures e.g., of advanced driver assistance systems; of training programs; of new road design treatments; of new traffic signalling arrangements; of new Police enforcement regimes; of new laws; etc.”

The importance of the NDS paradigm in overcoming the limitations of traditional methods of data collection and analysis in road safety is now well recognised by the international research community. The US, for example, has undertaken several large-scale NDS projects. The first was the seminal “100-car naturalistic driving study” (Klauer et al., 2006a), which explored factors leading to rear-end crashes, and the most recent (currently underway) is the US Strategic Highway Research Program Phase 2 (“SHRP 2”) NDS, which will deploy up to 3000 vehicles to explore and analyse a much wider range of road safety problems (Antin et al., 2011). Recently, Japan (Uchida et al., 2010), Europe (the EC-funded *UDRIVE* project), Canada and China have followed suit in ramping up their first large-scale NDS projects. To date, no large-scale studies of this kind and complexity have been undertaken in Australia.

Previous Naturalistic Driving Studies

To date, around 40 studies utilising the NDS approach have been undertaken worldwide (see Regan et al., 2012 for a review). Most have been small-scale studies. Several research issues have been examined, including factors leading to rear-end crashes (e.g. Klauer et al., 2006a); skill development in young drivers (Prato et al., 2010); skill loss in older drivers (e.g. Blanchard et al., 2010); young novice driver crash and incident types (e.g. Lee et al., 2011); distraction and inattention (e.g. Klauer et al., 2006b; Olsen et al., 2009; Koppel et al., 2011); fatigue (e.g. Hanowski et al., 2009); behaviour of drivers with dementia (Silverstein et al., 2009); interactions between light and heavy vehicle drivers (Hanowski et al., 2007); use of recorded data as feedback to improve driver safety (e.g. Toledo et al., 2008); understanding driver interactions with new vehicle safety technologies (e.g. Sayer et al., 2007); and lane changing behaviour (e.g. Lee et al., 2004).

While previous NDS projects have yielded some valuable insights into driver and road user behaviour in general, their applicability to the Australian context is questionable for several reasons. First, they have not yet explored many of the high priority, and intractable, road safety problems identified in the Australian National Road Safety Strategy (ATC, 2011). Speed choice and vulnerable road user interactions, in different situations, and in urban versus regional areas, are good examples (ATC, 2011). Second, it is not clear how well the findings translate to Australian conditions. Differences in cultural and societal norms, road laws, enforcement strategies, vehicle fleets, road environments, distances travelled, environmental conditions and mix of road users may threaten the transferability of data across countries. Finally, much data from NDS projects undertaken overseas (especially video data) are not accessible to Australian researchers for analysis, for ethical, commercial and other reasons (Regan et al., 2012).

The Australian 400-Car Naturalistic Driving Study

Project Partners

The Australian 400-car NDS, described in this paper, will be the first large-scale NDS undertaken in Australia. Led by the Transport and Road Safety (TARS) Research group, at the University of New South Wales, in Sydney, Australia, it will be an international project that brings together researchers from four leading universities in Australia – the University of NSW (via TARS), Monash University (via the Monash University Accident Research Centre), the University of Adelaide (via the Centre for Automotive Safety Research) and the Queensland University of Technology (via the Centre for Accident Research and Road Safety - Queensland) - and a leading transportation safety research institute in the United States (The Virginia Tech Transportation Institute; VTTI). The project also brings together several prominent road safety-related stakeholders from government and industry in Australia: the Centre for Road Safety at Transport for NSW; VicRoads; the Transport Accident Commission in Victoria; the Motor Accident Commission in South Australia; the Office of Road Safety, Government of Western Australia; and NRMA Motoring and Services Limited.

Aims

The overall aim of this project is to use the NDS method to understand what people actually do when they drive their cars in normal and safety-critical situations. The project will create a national facility in Australia, like that at VTTI in the USA, which can be used to support the running of NDS in Australia. The facility will be used over the coming decade to run multiple naturalistic driving studies in a range of vehicles (cars, trucks, and motorcycles) to answer a range of research questions, the first of these being the Australian 400-car NDS. The main outcomes of the first project (see below) will be the formulation, based on an in-depth analysis of driver and road user behaviour, of new and improved evidence-based countermeasures for improving intersection safety, reducing speed-related crashes, reducing fatal and serious injury crashes involving vulnerable road users (pedestrians, cyclists and motorcyclists), reducing crashes involving fatigue and inattention, and improving the design and performance of collision warning and other intelligent vehicle safety technologies.

Research Questions

While the NDS method can be used to answer an enormous range of research questions, the

Australian 400-car NDS will focus on seven key research themes, described below. These were selected jointly by the research, government and industry partners based on their significance for road safety, the lack of existing research on the sub-topics within each theme, their long term potential for program and policy development, value for money and feasibility to investigate. The seven research themes ranked most highly against these criteria are listed below.

Eight overarching research questions will guide the exploration of each of these research themes: (1) What collision risks are drivers exposed to?; (2) What collision risks do drivers expose themselves and others to?; (3) By how much do these exposures increase collision risk to them and to others?; (4) What are the factors that protect drivers against crash and injury?; (5) How do drivers modify their behaviour to adapt to conditions of increased and potential collision risk?; (6) How do people drive normally to minimise collision risk and avoid crashes?; (7) What are the factors that cause and contribute to crashes and safety-critical events?; (8) What is the relationship between driver personal characteristics, driving behaviour and crash risk?

The research themes are:

- ***Safety at intersections*** – the majority (around 32%) of all Australian serious casualty road crashes occur at intersections (ATC, 2011). Key issues to investigate include: crash and critical incident causation; speed choice versus speed limit at intersections; compliance with signage and signal phasing; gap selection when turning right across traffic; interactions with vulnerable road users; and situations in which crossing behaviour increases crash risk, and by how much.
- ***Speed choice*** – illegal and inappropriate speed are contributing factors in 34% of fatal crashes in Australia (ATC, 2011); key issues to investigate include: compliance with speed limit value and time of week; speed choice and lane width; interaction with traffic calming devices; gap acceptance during overtaking manoeuvres; speed choice versus traffic activity; speed choice versus level of pedestrian activity; speed choice versus type and age of vehicle; and situations in which speeding increases crash risk, and by how much.
- ***Interactions with vulnerable road users*** – vulnerable road users (pedestrians, cyclists and motorcyclists) account for 32% of deaths and 46% of serious injuries on Australian roads (ATC, 2011). Key issues to investigate in the NDS project include: crash and critical incident causation factors; gap and speed choice when passing cyclists; and anticipating pedestrian behaviours at crossings (lane and speed choice; when turning right and left).
- ***Fatigue*** – around 25% of fatal crashes in Australia involve fatigue as a contributing factor (ATC, 2011). Key issues to investigate include: level of fatigue as a function of distance travelled and time of travel; degree of fatigue versus type of driving (urban versus rural); analysis of the effects of fatigue on driving performance; and the situations in which fatigue increases crash risk, and by how much.
- ***Distraction and inattention*** - estimates from the US (which currently collects the most accurate data) suggest that distraction is a contributing factor in around 20% of all crashes (Gordon & Regan, 2012; Regan et al., 2009, 2011, 2012, Craft & Preslopsky, 2012). Unlike other factors such as speed, which can be calculated based on skid marks, tyre wear and level of impact, there is currently no post-crash method available to ascertain the involvement of distraction (and fatigue) in a crash, aside from often unreliable eyewitness accounts. Key issues to investigate include: measures of exposure for hand-held versus hands-free phone use; measures of exposure for adjusting console dials for radio, CD etc.; measures of exposure for texting and scrolling; adaptive behaviours when distracted; and situations in which distraction and other forms of inattention increase crash risk, and by how much.
- ***Crashes and near-crashes*** – around 250,000 crashes occur annually in Australia, many unreported. The proposed study will analyse all crashes and safety-critical events that occur in the NDS dataset.

- ***Interactions with intelligent transport systems (ITS)*** – collision warning and crash avoidance technologies known as ITS have enormous potential to prevent fatalities and serious injuries (Regan et al., 2006). Key issues to investigate include: driver response to variable message signs and electronic variable speed limit signs, and to over-speed and collision warnings.

The findings from the project will directly support injury countermeasure development activities for five of the nine major strategic challenges highlighted in Australia's National Road Safety Strategy for 2011-2020 (ATC, 2011). These include intersection crashes, vulnerable road users, speeding and rural driving.

Method

Design

The project will commence in the last half of 2013, with the first vehicles to be deployed in early 2014. A sample of 400 volunteer ordinary licensed drivers, 200 from NSW and 200 from Victoria, will be recruited. They will be broadly representative of Australian licenced drivers. Each driver will have their own car instrumented with data recording equipment for 6 months, allowing the researchers to record continuously their driving behaviours and those of others with whom they interact. The study will proceed in 6 phases, described below.

Previous NDS projects in the US have deployed instrumented vehicles for a year or more. This is because of major fluctuations in weather (e.g. snow) over the four seasons which have had to be controlled for in interpreting the data. In addition, the US studies have focused primarily on crashes and safety critical events, and hence a one-year deployment was necessary to yield enough events. In our study, major weather fluctuations will not be an issue in NSW and Victoria, the focus is not primarily on crashes, and the research themes of interest (described above) are ones involving behaviours that occur often and regularly. We believe on this basis that a 6-month deployment period is adequate.

The 100-car NDS in the US (Klauer et al., 2006a) yielded a dataset containing 69 crashes, 761 near-crashes and 8295 incidents. The objective in sample size selection for our study is to obtain a sufficiently large sample of incidents relevant to each of the research themes. As crashes will be the least frequent outcome of interest, the sample size of 400 was based on obtaining a sufficient sample of crashes. As crashes in this study will include minor and non-police-reported crashes, it is not possible to accurately predict the numbers expected to occur during the 6 month study period. The abovementioned US NDS study by Klauer et al. (2006a) provides the best basis for estimating crash frequency for the Australian NDS. This study found that 35% of drivers had at least one crash (including non-police reported crashes) and 14% had more than one crash (accounting for an additional approximately 50% of crashes). As all types of incident occurred evenly across the 12 month period, it is expected that the Australian study will result in 17.5% or 70 drivers having at least one crash in 6 months and there being around 100 crashes in total. As 78% of crashes in the US NDS involved inattention, we estimate that our study will yield at least 78 inattention-related crashes involving individual drivers. With around 12 % of these involving drowsiness in the US NDS we can expect around 9 drowsiness-related crashes only. For this research theme, near-crashes and incidents will be used as outcomes. Based on the ratios of near- crashes to crashes and incidents to crashes (11 times and 120 times, respectively) found in the US NDS, we expect at least 770 near crashes and at least 8,400 incidents, thus making it possible to describe less frequent road safety issues like drowsiness. Using the US NDS to calculate expected crash rates has problems as there may be differences in the nature of road exposure between the US and Australian NDS samples so affecting likely crash rates. However individuals within each country will also differ in the amount of exposure to the road environment they obtain, which may account for more variation than between-country differences.

The instrumented vehicles will be deployed in NSW and Victoria only. This reduces the logistic complexity of the project and saves money, while still providing an opportunity for interstate comparisons in driving behaviours. The vehicles will be deployed in one major city in each state (Sydney and Melbourne) and one selected rural location (probably Wagga Wagga and Bendigo).

Equipment

Almost all equipment required for the study will be purchased from VTTI, using funds obtained by the research team from a successful Australian Research Council (ARC) Linkage, Infrastructure, Equipment and Facilities (LIEF) grant. Major equipment items include 100 Data Acquisition System (DAS) units, 100 Mobileye units, computer storage hardware and software, data transfer computer servers, and data installation equipment and tools. The 100 DAS and Mobileye units will be rotated between 400 vehicles over a two-year period. Funding to run the study, and for some additional equipment needed, will be provided through a separate, successful, ARC Linkage grant. VTTI will also provide under contract to the University of NSW a range of essential expert services and advice as outlined below including installation training, data management and processing, technical support and troubleshooting, and quality control.

Project Phases

Phase 1: ***Driver Selection and Recruitment (Months 2-18)***

Selection. 200 drivers will be recruited from each of NSW and Victoria with equal numbers from city and country locations in each state. Drivers will a have valid Australian licence (provisional or full licence) and be aged between 17 and 70 years. The sample will be half male and half female. Drivers older than 70 years, the minimum age at which drivers are assessed for fitness to drive in most states, will be excluded as they are difficult to recruit, and drive relatively few kilometres per year. Participants must drive at least 5 days a week to collect sufficient data. Passenger cars will be the focus of the study, as they account for around 75% of all vehicles on the road. Vehicles to be instrumented will include a representative mix of sedans, coupes, hatchbacks, station wagons and sport utility vehicles (SUVs). Other vehicle types (e.g. trucks, motorcycles) will be instrumented and deployed in separate, follow-on, studies using slightly modified DAS units. Vehicle makes and models from 2002 onwards will be targeted, as these are more mechanically reliable, are more likely to enable access to vehicle network (CAN) data, and can be fitted with both the VTTI DAS and Mobileye systems (see below). It is recognised, however, that using vehicles less than 12 years old will create some bias in the results - by, for example, excluding from the sample older vehicles that might be used by younger drivers, drivers of lower economic means, and other driver sub-groups. The extent to which exclusion of vehicle makes and models prior to 2002 might limit the generalizability of the results is an issue currently under consideration by the research team.

Recruitment. Recruitment strategies found to be efficient in previous NDS projects overseas (Antin et al., 2011) will be used. Interested parties will log onto a custom-developed website and complete an expression of interest questionnaire that specifies the inclusion and exclusion criteria. Those meeting the criteria will be selected.

Phase 2: ***DAS installation in Vehicles and Driver Induction and Assessment (Months 2-24)***

DAS Installation. A dedicated Data Installation Site (DIS) will be created at each data collection site (city and country) in each State. Each DIS will comprise a garage with equipment necessary to install and de-install the DAS units. DAS installers (two at each Data Collection Site) will be trained in the US by VTTI. During installation, the DAS and Mobileye units will be fitted, aligned, calibrated, and checked to ensure that everything is operable.

The process for installing and de-installing DAS units will be the same in NSW and Victoria. For example, in NSW 50 units will be installed in Sydney and then another 50 in Wagga Wagga. Six months later de-installation of the first 50 Sydney units will commence in parallel with installation of the next wave of 50 units in Sydney. Immediately after Sydney, de-installation of the first 50 units and their re-installation in Wagga Wagga will occur. This will yield slightly more months of data for Wagga Wagga. 15 months after commencement of installation, de-installation of the final 50 units will occur in Sydney and then Wagga Wagga. This will yield data for 100 vehicles in Sydney and 100 in Wagga Wagga. This entire process will be mirrored in parallel in Victoria, i.e. Melbourne and Bendigo.

Induction. Participants will be inducted at the Data Installation Site (city or country) closest to them in each State. Written consent to participate will be sought from each driver. Basic demographic data will be collected including gender, date of birth, educational level, job type, working status, age of licensure, etc. In addition, vehicle inventory data will be collected (vehicle identification number, make, model, year, safety features, infotainment features, etc.).

Assessment. Each driver will be assessed on several dimensions relevant to driving (Antin et al., 2011), to allow us to explain variability in the driving data collected, and determine what dimensions predict observed driving behaviour. The dimensions to be assessed will be the same as those being assessed in the US 3000-car SHRP 2 NDS project, including assessments of higher cognitive function, visual perception, physical and psychomotor abilities, personality factors (e.g. propensity to take risks), sleep-related factors, medicines and medical conditions, and driving knowledge and history. Tests will be administered online and face-to-face, while vehicles are being instrumented.

Phase 3: ***Data Collection (Months 2 - 24)***

The DAS units were chosen to enable the researchers to collect all data required to answer the research questions of interest in this study. Each DAS unit comprises 3 main components: the *head unit, main unit* and a *forward radar assembly*. The ***head unit*** comprises: 3 low-light video cameras that record views of the forward roadway (in colour), driver face, and driver pedal and instrument cluster interactions; a still camera that takes images of passengers; a microphone for voice recordings; a cabin alcohol sensor; an incident pushbutton (for driver to record a 30 sec. verbal description of any incident of note); an illuminance sensor (to adjust for ambient lighting levels); and a GPS (to record latitude, longitude, elevation, time and speed). A separate camera will view the area behind the host vehicle. A Mobileye sensor will be co-located with the DAS head unit. Mobileye is a camera-based, multi-purpose, sensor that detects, automatically, in real-time, unintentional lane deviations, unsafe following distances, imminent collisions with a vehicle, motorcycle, cyclist or pedestrian, and detects speed limit signs.

The ***main unit***, located in the boot, hosts computer functions that coordinate sensor inputs, communications and data storage. It contains a 128 GB solid-state hard drive with capacity for 4-6 months of continuously recorded driving data, accelerometers (forward/reverse, right/left and up/down), yaw rate sensors (for crash and near-crash detection) and a mobile phone module. A forward radar will collect data on relative speed and position of forward traffic. Several machine-vision software applications will be incorporated into the DAS to enable automatic extraction from video footage of vehicle lane position, driver head position (from which can be extrapolated eye gaze direction) and driver identification. The information from the vehicle Car Network (CAN) will also be available for some cars.

Around 60 TB of data will be collected in total. The Machine-to-Machine (M2M) mobile telephone link set up between each DAS unit and the central computer located at UNSW will assess remotely the “health” status of each DAS unit, update DAS software, detect critical incidents and notify researchers when the hard drive is almost full. The DAS units were chosen because they are highly reliable. However, where problems cannot be rectified online, vehicles will be brought to the closest DIS for maintenance, with the consent of the driver. At the end of the study, crash-involved drivers will be interviewed by telephone to verify, through subjective feedback, the objective crash data collected. Every 4 to 6 months the near-full DAS hard disk, containing the continuous data, will be removed manually from the host vehicle and swapped for a fresh one (unless the participant has finished the study).

Phase 4. ***Data Management (Months 8 – 34)***

A *Data Management Office* (DMO) at two of the participating Australian universities (UNSW and Monash University) will securely house the data transfer and storage equipment: a “Drive Bay” (which reads the raw, encrypted, data recorded by each DAS hard drive and transfers it to a “Staging Server”); a “Staging Server” (which stores temporarily the data); a high bandwidth internet connection (for transfer of the raw data to a Central Staging Server located at UNSW); and USB drives (which are transported to UNSW, if there is data loss during internet data transfer). In addition, the DMO at UNSW will have a machine-to-machine (M2M) mobile telephone link to enable remote communication with each DAS unit in the field. The raw data from vehicles will be sent to a central computer at UNSW. It will then be sent via the Internet to the Virginia Tech Data Centre in the US, where it will be pre-processed by VTTI into a form ready for analysis by the Australian research teams; and securely stored in a database. To prepare the data for storage and analysis, it needs to be de-encrypted, synchronised (as not all data are collected at the same rate), and filtered (to eliminate outliers, etc.). This will be done by VTTI under contract to UNSW. To access the analysis-ready data, each Australian research team (NSW, VIC, QLD, SA) will generate data queries (using Client software; installed on their own computers), that are transmitted to the Data Centre at Virginia Tech via an Internet portal. The results will then be transmitted back to the researchers using the same Client software. The data pre-processed by VTTI will be downloaded regularly and stored securely in a Database Server at UNSW.

Phase 5. ***Data Analysis (Months 14 - 34)***

Data analysis is a multi-step process. *Data pre-processing* activities were described above. During *data processing*, developed algorithms will convert pre-processed data into “derived measures” such as time-to-collision (derived from radar and Mobileye data), average speed (derived from GPS data), minimum headway (derived from Mobileye data), lane excursions (derived from both DAS and Mobileye); etc. During *data enrichment*, the measures and derived measures will be combined with external data sources (e.g. geospatial information) to create new variables (e.g. speed limit, intersection, road lane width, pedestrian crossing, etc.) necessary for data analysis.

During *data reduction*, events and situations relevant to the research questions (e.g. overtaking, operating GPS, turning right, passing cyclist, passing through intersection, looking off road, etc.) will need to be detected in the data and coded. Where detection cannot be automated, coding will be manual, requiring multiple coders to search through video footage. A coding scheme, and annotation tools, will be developed for this purpose. Where possible, special algorithms (data “triggers) will be developed and calibrated for automatic search and extraction of safety-critical events (e.g. “hard braking” defined by -0.55 g longitudinal deceleration), critical events (e.g., turning right; overtaking) and baseline events (e.g., use of mobile phone). Triggered events will be validated by video-based verification. Coder training and quality control protocols (Klauer et al., 2011) will be developed and implemented.

Various *data analysis* techniques will be used to answer the eight overarching research questions for the research themes outlined above. *Exposure analyses* will identify and quantify the risks to which drivers and other road users are exposed. *Risk analyses* will derive measures (e.g. odd ratios) of increased or decreased crash risk associated with driver and road user exposure to these risks. *Data mining* techniques will be used to discover recurring patterns in driving behaviour associated with safety outcomes. *Descriptive analyses* will focus on the following: characterisation of the way in which drivers and other road users modify their behaviour to adapt to conditions of increased risk; characterisation of normal driving behaviours that minimise collision risk and avoid crashes; and identification and classification of factors that cause and contribute to crashes and safety-critical events. *Correlational analyses* will identify and quantify the relationship between driver characteristics, driving behaviour and safety outcomes.

Phase 6. ***Reporting (Months 35 - 36)***

A final project report will be provided to the four Industry Partners. Peer-reviewed papers and presentations at selected conferences will be produced throughout the project. Results will be presented regularly to Stakeholder Advisory Committee meetings held twice a year. A one-day workshop, convened by UNSW, will be held at the conclusion of the study to convey the key findings and recommendations, and discuss the lessons learned; to which the partners, key industry bodies, and state, national and relevant international agencies will be invited.

Governance Structure

For a large-scale and complicated project of this kind, a clearly defined governance structure is essential. TARS will manage and coordinate the project. MUARC will manage the deployment of vehicles in Victoria. There will be three Governance groups. The ***Core Research Group***, chaired by TARS, and involving all researchers and representatives from Transport for NSW and VicRoads, will refine and implement the study design, analyse the data and disseminate the findings. The ***Stakeholder Advisory Group***, chaired by one of the government partners, will ensure that all stakeholder requirements are met and fed back to the CRG. The ***International Advisory Committee***, chaired by the Chair of the Stakeholder Advisory Group, will involve 2 international experts with NDS project experience and will peer-review key deliverables and provide specialist scientific advice to the CRG and SAG. An ***NDS Governance Group*** (NDSGG) will oversee further use of the DAS and Mobileye units and data in future studies. The NDSGG will be chaired by TARS and involve representatives of each university and the industry/government partners. All equipment and data collected will be owned by UNSW. Exclusive licences will be granted to each Australian university to access and use the equipment and the data during the study for research purposes. Each Australian university partner, and VTTI, will have a copy of the raw dataset at the end of the project. After the project is over, each university partner, and through them each government/industry partner, will continue to have access to the data, to undertake further, separately-funded, analyses. Arrangements for access to the data by others is yet to be decided.

Conclusions

The NDS method overcomes a range of limitations associated with traditional approaches to data collection and analysis. There are, nevertheless, some potential methodological and other limitations associated with the NDS method: they are very resource demanding in terms of sample recruitment, data gathering, data storage and data analysis (Backer-Grondahl et al., 2009); driver behaviour may be influenced by knowledge of the presence of cameras and other sensors; crashes are rare events and thus very large sample sizes are needed to yield sufficient crash events; and, although large-scale studies may yield thousands of near-crash events, the validity of near-crash events as surrogates for crash events remains unresolved.

Despite the enormous range of issues that might be explored in NDS studies, relatively few have been explored, and many road safety issues of importance in Australia remain to be explored using this method. We have outlined in this paper what we see as the advantages in running naturalistic driving studies in Australia, and more generally the advantages they could offer in complementing existing techniques for gathering data to support countermeasure development. The Australian 400-car NDS, even on its own, will provide a massive “living” database of information that can be interrogated for many years to improve countermeasure development and improve Australia’s road safety performance relative to other OECD countries. The data collection and analysis facility developed as part of this project will be used over the coming decade to run further naturalistic driving studies in Australia in a range of vehicles (cars, trucks, and motorcycles) to answer a broader range of research questions.

Acknowledgements

The authors would like to acknowledge the support of the government and industry partners, listed in this paper, who pledged financial and in-kind support for the 400-car Australian NDS. We thank also the following colleagues who supported the authors in securing a successful ARC LIEF grant that will fund the purchase of most of the equipment for this study: Prof. Mark Stevenson, from MUARC and Associate Prof. Paul Salmon, from the University of the Sunshine Coast.

References

Antin, J., Lee, S., Hankey, J. & Dingus, T. (2011). Design of the in-vehicle driving behaviour and crash risk study. Transportation Research Board: Washington DC, USA.

ATC, Australian Transport Council (2011). National Road Safety Strategy 2011-2020. Canberra, ACT: ATC

ATSB (2002). *International Road Safety Comparisons 2002*. Canberra, ACT: Australian Transport Safety Bureau.

ATSB (2005). *International Road Safety Comparisons 2005*. Canberra, ACT: Australian Transport Safety Bureau.

Backer-Grondahl, A., Phillips, R., Sagberg, F., Touliou, K., & Gatscha, M. (2009). Topics and applications of previous and current naturalistic driving studies. Deliverable D1.1. PROLOGUE project. Brussels, Belgium: European Commission.

Blanchard, R. A., Myers, A. M., & Porter, M. M. (2010). Correspondence between self-reported and objective measures of driving exposure and patterns in older drivers. Accident Analysis and Prevention, 42, 523-529.

Bureau of Infrastructure, Transport and Regional Economics (BITRE), 2009. *Cost of road crashes in Australia 2006,* Report 118. Canberra.

Craft, R.H. and Preslopsky, B., (2012). Distraction and Inattention: Top Crash Causes in the USA (Chapter 8). In Regan M.A., Lee J.D. and Victor T.W., Driver Distraction and Inattention: Advances in Research and Countermeasures, Volume 1, Surrey, England, Ashgate, ISBN 978140942584 – in press. Date of acceptance: 6 June 2011.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. et al. (2006). *The 100-car naturalistic driving study: Phase II – Results of the 100-car field experiment* (Report no. DOT HS 810 593). National Highway Traffic Safety Administration (NHTSA), Washington, DC.

Gordon C. and Regan M.A., (2012). Driver Distraction and Inattention and Their Role in Crashes and Safety-Critical Events (Chapter 9). In Regan M.A., Lee J.D. and Victor T.W. (Eds) Driver Distraction and Inattention: Advances in Research and Countermeasures, Volume 1, Surrey, England, Ashgate, ISBN 978140942584 – in press. Date of acceptance: 6 June 2011.

Hanowski, R. J., Hickman, J. S., Olson, R. L., & Bocanegra, J. (2009). Evaluating the 2003 revised hours-of-service regulations for truck drivers: The impact of time-on-task on critical incident risk. Accident Analysis and Prevention, 41, 268-275.

Hanowski, R. J., Hickman, J. S., Wierwille, W. W., & Keisler, A. (2007). A descriptive analysis of light vehicle–heavy vehicle interactions using in situ driving data. Accident Analysis and Prevention, 39, 169-179.

International Transport Forum (2011). IRTAD Road Safety Annual Report 2010, Geneva, Switzerland: OECD.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., & Ramsey, D.J. (2006A). The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data. Report No. DOT HS 810 594, NHTSA: Washington, D.C.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., & Ramsey, D.J. (2006B). The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car NDS data. Report DOT HS 810 594, Nat. Highway Traffic Safety Admin.: Washington, D.C: NHTSA

Klauer, S.G., Perez, M., & McClafferty, J. (2011). Naturalistic driving studies and data coding and analysis techniques. In B.E Porter (Ed). Handbook of traffic psychology. London: Academic Press.

Koppel, S., Charlton**,** J**.**, Kopinathan, C., & Taranto, D., (2011). Are child occupants a significant source of driving distraction? Accident Analysis and Prevention, Volume 43, Issue 3, May 2011, 1236–1244.

Lee, S. E., Olsen, E. C. B., & Wierwille, W. W. (2004). A comprehensive examination of naturalistic lane-changes (Report No. DOT HS 809 702). Washington, DC: National Highway Traffic Safety Administration (NHTSA)

Lee, S. E., Simons-Morton, B. G., Klauer, S. E., Ouimet, M. C., & Dingus, T. A. (2011). Naturalistic assessment of novice teenage crash experience. *Accident Analysis and Prevention, 43*, 1472-1479.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations. Report No: FMCSA-RRR-09-042. U.S. Department of Transportation, Federal Motor Carrier Safety Administration. Washington DC: FMCSAPrato C.G., Toledo T., Lottan T., & Taubman – Ben-Ari., O. (2010). Modelling the behaviour of novice young drivers during the first year after licensure, Accident Analysis and Prevention, 42, 2, 480-486.

Regan, M.A., Hallet, C, & Gordon, C (2011). Driver Distraction and Driver Inattention: Definition, Relationship, and Taxonomy. Accident Analysis and Prevention, 43, pp. 1771-1781.

Regan M.A., Lee J.D. and Victor T.W., (2012) Driver Distraction and Inattention: Advances in Research and Countermeasures, Volume 1, Surrey, England, Ashgate, ISBN 978140942584 – in press (Hardcover). Date of acceptance: 6 June 2011.

Regan M., Lee J.D., and Young K., 2009. Driver Distraction: Theory, Effects and Mitigation. CRC Press, Taylor and Francis, Boca Raton, Florida (34 Chapters), ISBN 13-978-0-8493-7426-5 (Hardcover).

Regan, M., Triggs, T., Young, K., Tomasevic, N., Mitsopoulos, E., Stephan, K., & Tingvall, C. (2006). On-Road Evaluation of Intelligent Speed Adaptation, Following Distance Warning and Seatbelt Reminder Systems: Final Results of the Australian TAC SafeCar Project. Volume 1: Report. Monash University Accident Research Centre Report 253. MUARC: Melbourne, Australia.

Regan M.A., Williamson A., Grzebieta R**.**, Tao L., (2012). Naturalistic Driving Studies: Literature Review and Planning for the Australian NDS, Proc. Australasian College of Road Safety Conference – A Safe System: Expanding the reach!", Sydney.

Sayer, J., Devonshire, J., & Flannagan, C. (2007). Naturalistic driving performance during secondary tasks. In: Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 9-12 July, Stevenson, Washington.

Silverstein, N. M., Eby, D. W., Molnar, L. J., LeBlanc, & D. J., & Adler, G. (2009). Monitoring drivers with dementia: An instrumented vehicle study. Alzheiemer's and Dementia, 5, 140-140.

Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behaviour. Transportation Research Part C: Emerging Technologies, 16, 320-331.

Uchida N, Kawakoshi M, Tagawa T, & Mochida T. (2010). An investigation of factors contributing to major crash types in Japan based on naturalistic driving data. IATSS Research, 34(1), 22-30.