EVALUATION FRAMEWORK FOR THE QUEENSLAND ROAD SAFETY STRATEGY 2015-2021

Final Report

December 2016

Angela Watson, Mark King, Amy Williamson, Ross Blackman, and Jason Edwards

The Centre for Accident Research & Road Safety – Queensland is a joint venture initiative of the Motor Accident Insurance Commission and Queensland University of Technology

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Queensland Road Safety Strategy 2015-2021 (the Strategy) represents a shift in the Queensland Government's approach to road safety management. A key aspect of this shift relates to internalising the attitudes, beliefs and values underlying the Safe System approach, assisted through the development and promotion of traffic safety culture (TSC). Some other elements of the current Strategy differ from those in the previous two strategies, including a greater focus on serious injuries as well as fatalities, and greater attention to distraction, inattention and drug driving. The current Strategy has also seen the introduction of "safe speeds" as a Safe System cornerstone, in addition to "safe roads and roadsides, safe vehicles and safe road users".

Progress towards achieving these broad objectives needs to be monitored through rigorous evaluation. However, the existing road safety strategy evaluation framework does not account for the need to measure a targeted cultural shift, nor monitor progress in relation to other new strategy elements and objectives. The Queensland Department of Transport and Main Roads (TMR) therefore recognised a need to review the existing framework to ensure that the framework is tailored specifically for comprehensive evaluation of the new strategy.

TMR commissioned the Centre for Accident Research and Road Safety – Queensland (CARRS-Q) to update Queensland's existing road safety strategy evaluation framework. The 2004-2011 and 2015-2021 Strategies were reviewed and compared to identify similarities and differences in terms of scope and objectives. The existing evaluation framework was also reviewed with these similarities and differences in mind. The following specific tasks were undertaken to achieve the project aims:

- Analysis of the existing evaluation framework in the context of the 2004-2011 and 2015-2021 Strategies
- Provision of recommendations to update the framework in light of the target, vision and objectives of the 2015-2021 Strategy
- Update of the framework in line with recommendations and the existing GOSPA model
- Development of data collection instrument/s required to implement the evaluation framework
- A "test run" of the framework using available data
- Provision of milestones for application of the framework

The existing evaluation framework employed a "GOSPA" model for strategy evaluation, referring to the multi-tiered framework for evaluation stratified in terms of **G**oals, **O**bjectives, **S**trategies, **P**rograms/Plans, and **A**ctions. Following a review of relevant literature by CARRS-Q for the current project, the GOSPA model was retained as the basis for the revised multi-tiered evaluation framework, comprising the **Global Assessment**

(First Tier) Model; the **Second Tier Model**; the **Third Tier Model**; and the **Supplementary Model**.

The **Global Assessment (First Tier) Model** relates to the broad goals and objectives of the Strategy. Specifically, the objectives of the 2015-2021 Queensland Road Safety Strategy are to:

- Reduce fatalities from 303 (average 2008-2010) to 200 or fewer by 2020
- Reduce hospitalised casualties from 6,670 (average 2008-2010) to 4,669 or fewer by 2020

It is proposed that the First Tier assessment in the existing framework remain in its entirety. Additionally, it is proposed at this level to also include monitoring of progress toward serious injury targets (using improved data and analysis) as well as fatalities.

The **Second Tier Model** relates to the specific strategies and programs of the Strategy and the first Action Plan. For Second Tier assessment, the existing framework included stratification by broad road user groups and issues (e.g., impaired driving, fatigue, pedestrians, older/younger road users, fleet-related). These analyses are retained in their entirety in the revised framework, but there are modifications and additional strategy components to be evaluated at this level. This includes progress in terms of community and stakeholder attitudes, beliefs, values and behaviours (conceptualised as traffic safety culture). Another additional inclusion proposed for this level employs remoteness indices for a finer level of analysis than is provided by the urban/rural split used in the existing framework. The existing framework also does not explicitly address strategy goals and objectives in terms of Safe System "cornerstones" and does not identify "safe speeds" as a Safe System element due to its absence from previous strategies. These issues would also be addressed in the revised framework at the Second Tier level.

The **Third Tier Model** deals with individual program elements and actions of the Strategy and associated Action Plans. It is suggested that the analyses recommended in the existing framework for this level be retained for the revised framework, including assessment of specific road safety program activities such as speed limit enforcement, RBT, other enforcement measures, mass media campaigns and changes in crash reporting. However, in light of the current Strategy and action plan, the revised framework for this tier should address additional specific issues and strategy elements, including for example; roadside drug testing; cycling-related regulations and enforcement; increased attention to distraction, related penalties and enforcement (e.g., mobile phone offences); social media activity and related themes; changes to Q-Ride.

The **Fourth** or **Supplementary Tier** is recommended for specific evaluation of major program components. The first three models introduced above measure the general association between measures of specific road safety program activities in a multivariate setting. However, for large complex road safety program elements, specific evaluation of

major elements is considered necessary because; (i) only specific evaluation can *reliably* establish cause and effect relationships between program element implementation and road trauma outcomes; and (ii) specific evaluation is often needed to establish the measure of road safety program operation that is best related to the outcomes achieved. This information is fed back into earlier model/s as a key input, and relies largely on the commissioning by TMR of evaluations of specific programs and legislative changes that support the goals and objectives of the Strategy.

The main body of the report details precisely why the GOSPA model is most appropriate, operationalisation of the model and specific methodologies applied, data collection instruments used, the scope and nature of anticipated results, and the limitations that have been identified. On the basis of a comprehensive review of the relevant literature, the concept of traffic safety culture and its relation to the Safe Systems approach is also discussed.

Finally, specific recommendations offered in light of the review of literature, the current Road Safety Strategy, and the previous evaluation framework include:

- Measures of traffic safety culture and systems thinking be included in the framework, and developments in related theory and measurement be monitored
- Hospitalisations (based on hospital data) be included (in addition to the current definition) as a measure of serious injury in trend modelling and evaluation
- Cyclists be included as a strata (grouping) for Second and Third Tier analysis
- Refinement of the rural and remote classification for stratification in the Second and Third Tier to match the Australian Standard Geographic Classification (ASGC) Remoteness Structure
- Distraction as a contributor to road crashes should have more emphasis in the Second and Third Tier assessment
- Improvements to the measures of distraction and its role in crashes be investigated
- Improvements to data to better distinguish between drug and drink driving and to better quantify their relative involvement in crashes be investigated
- Opportunities for data linkage and other data improvement strategies be monitored

1 INTRODUCTION

The Queensland Department of Transport and Main Roads (TMR) commissioned the Centre for Accident Research and Road Safety – Queensland (CARRS-Q) to update Queensland's existing road safety strategy evaluation framework.

1.1 Background

The Queensland Road Safety Strategy 2015-2021 represents a shift in the Queensland Government's approach to road safety in Queensland, seeking the development and strengthening of a traffic safety culture (TSC). Under the Safe System approach there has been increasing recognition of the need for stakeholders in all parts of the Safe System to internalise the attitudes, beliefs and values that underlie the Safe System, i.e. to undergo a cultural shift towards achievement of a traffic safety culture. While reliance on enforcement and regulation has been very effective, long term sustainable change that is endorsed by all stakeholders can only be achieved through voluntary commitment to the principles and practices of a Safe System.

However, the current road safety strategy evaluation framework does not take account of the need to measure this targeted shift in culture. It is a comprehensive and carefully constructed approach that focuses on the measurement of crash outcomes, designed to address the targets in the previous strategy. TMR recognises a need to review the framework for two purposes: (i) to ensure that the framework is tailored to address the evaluation of the new strategy; and (ii) to incorporate the monitoring and measurement of cultural change, i.e. traffic safety culture, among both the general population and major stakeholders. We suggest that a third purpose can be added, that relates to the inclusion of a serious injury target in national and State road safety strategies: given the lags in reporting of serious injury crashes and the varying levels of underreporting by crash type/mode (e.g. single vehicle bicycle injury crashes are mostly unreported to police) and degree of injury, there is a need to incorporate these challenges into the evaluation framework.

The existing evaluation framework was tested in a "dry run" using historical crash data, and this approach will be followed for the revised evaluation framework; however this will not address the measurement of traffic safety culture, which is typically measured at community level through surveys. Given the historical absence of such surveys in Queensland, it is proposed that Wave surveys and other reasonably representative surveys will be examined to determine whether they have used survey items similar to those which would be used in a traffic safety culture survey. At the stakeholder level there is likely to be little existing data to work from, however CARRS-Q will investigate available sources to determine if this is the case.

1.2 Project aims

The purpose of the evaluation was to update Queensland's existing road safety strategy evaluation framework. The following tasks were undertaken to achieve the project aims:

- Analysis of the existing evaluation framework in the context of the 2004-2011 and 2015-2021 Strategies
- Provided recommendations to update the framework in light of the target, vision and objectives of the 2015-2021 Strategy
- Updated the framework in line with recommendations and the existing GOSPA model
- Development of data collection instrument/s required to implement the evaluation framework
- Performed a "test run" using available data
- Provided milestones for application of the framework (i.e. after the implementation of every action plan)
- Prepared a report summarising project activities.

2 METHODOLOGY

This chapter outlines the approach taken to deliver the research services specified by TMR. The following breakdown of tasks was:

Task 1: Analyse the existing evaluation framework in the context of the 2004-2011 and 2015-2021 Strategies.

Task 2: Provide recommendations to update the framework in light of the target, vision and objectives of the 2015 – 2021 Strategy.

Task 3: Update the framework in line with recommendations and the existing GOSPA model.

Task 4: Develop data collection instrument/s required to implement the evaluation framework

Task 5: Perform a "test run" using available data.

Task 6: Provide milestones for application of the framework (i.e. after the implementation of every action plan).

2.1 Analyse the existing evaluation framework in the context of the 2004-2011 and 2015-2021 Strategies

The 2004-2011 and 2015-2021 Strategies were reviewed and compared to identify similarities and differences in terms of scope and objectives. The existing evaluation framework was also reviewed with these similarities and differences in mind. This process determined the alignment between the objectives of the existing framework with those unique to the 2015-2021 Strategy. This allowed for the identification of the elements of the existing framework that are recommended to be maintained in the revised framework and the elements that need to be included that were not previously. The review process was conducted within the GOSPA[1] framework with assessments made as to how the aspects of the 2015-2021 Strategy are incorporated within this model. The systematic approach to the review is presented in Table 1.

[1] GOSPA refers to the framework for multi-tiered strategy evaluation in relation to Goals, Objectives, Strategies, Programs/Plans, and Actions, as fully described in Section 3 and summarised below in Table 1.

Table 1: Approach to analysis of Strategies and evaluation framework

GOSPA component	What is this for the 2015-2021 Strategy?	How does this align with the 2004-2011 Strategy?	How does (if at all) the existing evaluation framework address this?	What additional evaluative components can be added to address this?
Goal: General (idealistic) statement of the overall goal	e.g., reduce road trauma, expand our understanding of the "road toll", drive a fundamental change in the culture and attitude to road safety	Align completely/partially/no alignment	e.g., forecasting model for crashes	e.g., alternative measures of serious injury, measures of cultural change
Objectives: Specific (pragmatic) statements of the measurable objectives to reach the Goal	e.g., reduce fatalities to 200 or fewer by 2020; reduce hospitalised casualties to 4,669 or fewer by 2020	Align completely/partially/no alignment		
Strategies: General (idealistic) strategies to achieve each Objective	e.g., education and engagement; technology	Align completely/partially/no alignment	e.g., strata specific forecasting models	e.g., measures of cultural change
Programs/**P**lans: Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs	e.g., engage community, industry, and other stakeholders; plan and implement emerging infrastructure technologies to make road use safer	Align completely/partially/no alignment		
Actions: Actions undertaken in each program.	e.g., State-wide advertising and social media campaigns	Align completely/partially/no alignment	e.g., evaluation of specific programs	e.g., include alternative measures of serious injury

2.2 Provide recommendations to update the framework in light of the target, vision and objectives of the 2015 – 2021 Strategy

In order to inform the recommendations there was a review of the literature relating to cultural change, safety culture, attitudinal change, and serious injury measurement. Internet searches (e.g., Google, Google Scholar, Research Gate) were conducted as well as searches of the following databases:

- Australian Transport Index (ATRI)
- Proquest
- PsychINFO
- ScienceDirect
- TRIS Online (Transportation Research Information Services).

This review of the literature, as well as the expertise of the team, allowed for the incorporation of measures of "Cultural Shift" within the framework. The literature review focused on how the conceptualisation and measures of safety culture can be adapted to traffic safety culture for stakeholders and the community at large.

Data requirements and availability were also reviewed as part of this Task. This enabled recommendations to be made surrounding the specific data requirements for each component of the framework. Issues identified related to the relevance (the ability of the data to address the specific component), accessibility (whether access by TMR and/or the evaluation researchers is possible and if so how it can be accessed), timeliness (are there delays in data availability and what would be the impacts on the timing for action plan evaluation), and quality of the sources (completeness, reliability, and validity).

2.3 Update the framework in line with recommendations and the existing GOSPA model

The framework was updated within the GOSPA model based on the outcomes of Task 1 and Task 2. The framework describes the alignment between the GOSPA model and the current Strategy. It then describes the process of the development of the framework and, for each level of the GOSPA model, describes the specific methodologies that are recommended to be applied. The framework includes a description of available data sources as well as the proposed data collection instruments (see below for more detail). The framework describes and justifies all statistical analyses that are to be used. The issues and limitations are also included as well as a description of all extraneous variables and/or factors that will be measured and controlled for/considered.

2.4 Develop data collection instrument/s required to implement the evaluation framework

As part of the literature review (described above) best practice measures of traffic safety culture and attitudinal change were identified. These pre-existing measures have been mapped where possible to apply specifically to the current Strategy. Any previously validated instruments that are recommended for inclusion have been adapted to apply specifically to the Queensland context where possible. Recommendations for the next steps in survey development have been described.

2.5 Perform a "test run" using available data

Based on the evaluation framework, available data was used to demonstrate its use. To test-run the trend analysis aspect of the framework, structural time series models were fitted to police-reported fatalities and police-reported "hospitalisations". These models use data that was consistent in definition to that used by the previous framework. In addition, time series models were produced using hospital data for those traffic-related injuries admitted to hospital and traffic-related injuries considered to be high threat to life. See Section 7 for details.

3 ANALYSIS OF THE EXISTING EVALUATION FRAMEWORK IN THE CONTEXT OF THE 2004-2011 AND 2015-2021 STRATEGIES

3.1 Review of the existing framework

The road safety strategy evaluation framework developed by the Monash University Accident Research Centre (MUARC, 2009) was reviewed to assess its suitability and applicability in terms of evaluating the current Queensland Road Safety Strategy 2015 – 2021 (Queensland Transport and Main Roads, 2015). The MUARC framework was trialled on a previous Queensland Road Safety Strategy (1993 - 2003, revised in 1998), and is a modified version of the framework outlined by Cameron (1999), presented as Appendix A in MUARC (2009).

3.1.1 The GOSPA model

A review of relevant literature by MUARC (Section 2) found that the GOSPA model provided the most appropriate foundation for defining elements of the framework, facilitating a multilevel strategy evaluation linked to and structured in terms of Goals, Objectives, Strategies, Programs and Plans, and Actions. Importantly, the GOSPA model "relates evaluation to a strategic planning framework" (p. 20) such as used in the previous two Queensland road safety strategies: "to a large extent, the 1993-2003 and 2004-2011 strategies and action plans are structured in terms of an overall goal, objectives, strategies, programs/plans, actions and targets" (p. 20).

MUARC (2009) describes the framework as a multi-tiered modelling approach which had not then been previously used in road safety strategy evaluation. There are three main tiers in the model as described in this instance, plus a fourth "supplementary" tier which relies on external evaluations of specific actions and program activities. It is apparent that the fourth "supplementary" tier is necessary to feed back into the third tier assessment model and is therefore essential to the overall framework.

1. **Global or 1st Tier Assessment Model** - Performance regarding broad goals and objectives; measure reduction in overall trauma, against "that expected based on pre-strategy implementation trends"
2. **2nd Tier Assessment Model** - Performance regarding targets in strategies
3. **3rd Tier Assessment Model** - Relate trends to explicit measures
4. **Supplementary tier**.

MUARC (2009, p. vi – vii) describes the GOSPA framework as defining:

> *… a pyramid of increasing detail in defining the elements of a road safety strategy. The top of the pyramid defines the broad goals for which the strategy is aiming (Goals) whilst the next level down gives specific measurable targets (Objectives) against which the goals can be assessed. The Strategies area of*

the framework typically defines the target areas on which the road safety strategy will focus to achieve its goals along with local objectives within each target area that will jointly contribute to achieving the global objectives. Finally the Programs and Actions areas contain the specific details on the type of activities to be carried out in each target area and the amount of effort that will be applied to each activity.

With five components in the three-tier GOSPA framework, Tier 1 covers the strategy Goals and Objectives, Tier 2 concerns Strategies, while Tier 3 covers the specific Programs, Plans and Actions of the overall strategy. The "pyramid" model/framework can be diagrammatically represented as in Figure 1, below.

Figure 1: GOSPA framework

First Tier/Global Assessment Model

Although it can be formulated in different ways, the main purpose of the global assessment model is to measure the effectiveness of the overall strategy in reducing road trauma. Forecasts will be modelled at this level to estimate the future trends expected in the presence or absence of various road safety programs and countermeasures, based on pre-strategy trends. Strategy effectiveness can be measured at this level in terms of different severity levels (i.e., fatal crash risk, serious casualty crash risk and all injury crash risk) provided sufficient data are available, taking into account the broader goals, objectives and directions of the strategy. Contaminating and confounding influences (i.e., socio-economic) were removed from global assessment model in the MUARC example to more accurately assess the effect of initiatives.

MUARC found that forecasting beyond the first two years of a strategy (or Action Plan) may be somewhat unreliable due to the "typically wide confidence limits" produced by analysis at this stage. As a result, outcomes may be measured incrementally at regular intervals for the most reliable assessment (p. 99):

> *…there may be a need to re-calibrate the forecasts at regular intervals, say every 2 years to correspond with the development of each new action plan, based on a longer data period.*

Second Tier Assessment Model

The Second Tier Assessment model is a more detailed assessment which considers strategy effects in terms of specific groups of road users, circumstances and conditions or other variables, defined as strata. In the example given for Second Tier Assessment, MUARC includes the following groups, or strata (p. 50):

- *Alcohol and drug-driving (modelled during high alcohol times of the week)*
- *Speed-related crashes and speeding drivers*
- *Fatigued drivers*
- *Young adult inexperienced drivers*
- *Older drivers*
- *Fatal and serious crashes in rural Queensland*
- *Pedestrians, including intoxicated pedestrians*
- *Unrestrained vehicle occupants*
- *Unlicensed drivers and riders*
- *Motorcycle riders*
- *Indigenous road users*
- *Roads with poor crash records*
- *Fleet-related crashes.*

It may be worth noting here that these strata do not include cyclists or distraction and inattention as target areas.

The issue of reliability of analyses and forecasting beyond the first two years of a strategy, as mentioned above for the First Tier, is also present at the Second Tier level.

Third Tier Assessment Model

The strata targeted in the Second Tier Assessment, as presented above, are also targeted in the Third Tier model, but the Third Tier model also includes specific measures under different activity areas as covariates.

Description of the evaluation of the Queensland Road Safety Initiatives Package (RSIP) - related to "Fatal Four" enforcement, associated media and education - was provided as an example of the type of modelling performed in Third Tier analyses. "RSIP can be thought

of as representing a mini road safety strategy" (p.72). There were 32 strata of analysis in total. Analysis was stratified by Police Region (n=8), road type (n=2) and high/low alcohol hours (n=2). Relevant data were also collapsed into three severity levels (fatal & serious injury; medical treatment injury; other injury & no injury).

Poisson regression (modified model) was used for the measuring of level and significance of association between RSIP measures and crash outcomes. A structural time series model could also have been used for this analysis, but was considered less straightforward to apply "in an integrated manner to a stratified analysis" (p.76).

Analysis accounted for elements of the RSIP that were already in place, but may have been implemented at a lower intensity.

Supplementary Tier

The Fourth or Supplementary Tier incorporates specific program activities and refers to the use of existing methods. As this level of evaluation is "outside the scope" of the main framework, it relies on dedicated separate evaluations of specific programs (e.g. Random Road Watch, Mobile Speed Camera Program, 50km/h default urban speed limit, p. 83) as commissioned by the relevant authority.

As well as crash data, valuation relies on numerous before/after surveys, for example, quasi exposure measures, free-flow speeds, bicycle helmet and seatbelt usage rates, enforcement and offence data (e.g., BAC testing prevalence and results), road user attitudes etc. New Zealand surveys are cited as examples of the type that are suitable to use. Repeated surveys are to be conducted at the same places and times under similar weather conditions etc.

Additional to surveys, assessing the impact of road safety advertising requires regular collection of advertising data for each campaign, separable by theme (speed, drink driving etc.) and style (i.e., emotive, educational etc.).

Limitations

It is noted that trends measured over short timeframes (i.e., 2, 3, 5 years) may not be at all reliable. There is no avoiding this when evaluating performance of a strategy over the short and medium terms (as is required to inform 2-year Action Plans etc.).

The MUARC framework does not address strategy goals, objectives and outcomes in terms of "Safe System" cornerstones.

MUARC (2009, Table 5.2) refers to Vulnerable Road Users as a key road user group comprising rural road users, older road users and young/inexperienced drivers. While these road users may be described as being generally at greater risk of road traffic injury than the general population, the term Vulnerable Road Users commonly refers in the

literature to motorcyclists, pedestrians and bicycle riders. Although this is a relatively minor issue of terminology, it may be preferable to remain consistent where possible with common road safety phrases and terms.

Beyond stratification by the eight Queensland Police Regions, the geographic separation in the MUARC framework is limited to a distinction between "urban" and "rural" road environments, defined as less than 80km/h for urban and equal to or greater than 80km/h for rural. Such a distinction is arguably of limited value as it is not likely to reveal potentially important differences between regional, remote and very remote areas in terms of crash (and road user) characteristics and road safety program activities such as enforcement, advertising and engineering measures (among others). Analyses utilising remoteness indices such as those included in the Australian Standard Geographical Classification (ACGS) are likely to be useful for helping to overcome this limitation. Any further stratification based on remoteness should be weighed against any power issues of small numbers (see Section 3.3.6).

3.1.2 Other approaches

While the GOSPA model was chosen as able to provide the most appropriate framework in this instance, other approaches may also be adopted. MUARC's completion of numerous evaluations of various Australian road safety initiatives is offered as evidence that there is "no single best method" (p.12) for road safety program evaluation; all methods require assumptions and the design of the evaluation framework will be influenced by the program nature and the vision of stakeholders.

The GOSPA model was one of two main methods identified by MUARC for evaluating road safety strategies. The second method identified was the PIARC (World Road Association) procedure, described as a "four-step procedure for a systematic framework of evaluation" (p. 7). This is based on methods typically used for cost-benefit analysis, but also incorporates multi-criteria analysis to assess impacts that may be better evaluated in non-economic terms. The PIARC study/procedure is not described in detail, but it is noted that processes of project evaluation can be diverse, that single methodologies are often unacceptable, and that method selection depends substantially on the road safety issue in question as well as the data available. The four basic steps of the PIARC procedure include (p. 7):

a. *Defining gains and losses according to some set of objectives*
b. *Listing advantages and disadvantages*
c. *Measuring gains and losses in some unit or different units*
d. *Decisions are made on the basis of (i) to (iii) by using explicit weights of importance (e.g. monetary values) or implying such weights ex-post.*

A control chart technique was used in a previous evaluation of the (1993 – 2003) Queensland road safety strategy. In this technique, problems were identified and ranked, forming the rows of the chart, while options to address the problem areas form the chart columns (the MUARC document seems to have described these rows/columns in reverse).

The impacts of individual programs could be estimated, and reportedly also aggregated to provide an assessment of overall ("global"?) strategy effectiveness and performance. Results showed that (p.9):

> *…the aggregation of countermeasures method could realistically only account for 80% of the crash savings identified using the macro model which estimated the overall impact of the strategy (Leggett 1999). Furthermore, the aggregation method could not take into account any positive crossover effect one program may have had on another.*

These limitations are able to be overcome to some extent using the GOSPA model as described. The ability to account for the combined impact of multiple interventions and to also estimate their separate effects is crucial where more than one initiative is applied to address a specific road safety issue.

The Western Australian road safety strategy 2003-2007 is one of few that have been evaluated in terms of the effects of each program, using outcome measures (e.g., crash rates), intermediate measures (e.g., attitudes/behaviour) and process measures (e.g., intensity of enforcement). While the statistical methods (model/s) used are not described in MUARC (2009), the approach reportedly allowed regular and timely progress reports to be produced throughout the life of the strategy.

Some European road safety program evaluations are also summarised in Section 2 of MUARC (2009), including the ESCAPE and SUNflower projects. The SUNflower project sought to identify reasons for the strong road safety performance of Sweden, Netherlands and the UK relative to other European countries, and to inform further improvements. The specific methods used are not described, but as with other road safety program evaluations, it is reported that the impact of major initiatives can often be estimated with relative ease while the effectiveness of smaller initiatives may not be measurable. Outcome measures were limited to road traffic fatalities and fatality rates and did not address the problem of serious injuries. The ESCAPE project focused on common driver behavioural issues and the potential of enforcement measures to improve road safety outcomes. In its summary of the project, MUARC focused on issues around the selection of appropriate performance indicators, research design, data collection and analysis. Within the scope of the ESCAPE project, the performance indicators would relate to crash statistics, enforcement activity, and road user attitudes and behaviours. In terms of data analysis, as with other projects, the most appropriate method depends on the nature of the data and the specific questions asked.

Overall, the literature review by MUARC (2009) illustrates the importance of several key issues for developing an evaluation framework. Among these are the selection of the most appropriate performance indicators, the reliability, timeliness and scope of available data, and the potential to assess and estimate both the combined and separate impacts of multiple actions and initiatives.

3.2 Elements to be retained from existing framework

Imperfections in data collection and availability notwithstanding, a framework founded on the GOSPA model is a highly appropriate suite of methods to evaluate the current Queensland Road Safety Strategy.

3.3 Elements missing from existing framework

Some of the elements and emphases of the current Queensland Strategy differ from those in the previous two strategies. These include, in the current Strategy, a greater focus on tracking and addressing serious injuries as well as fatalities, greater attention to distraction, inattention and drug driving, and explicit efforts to drive a change in road safety attitudes and behaviours at a cultural level.

3.3.1 Safe System approach

The MUARC framework does not clearly or fully encapsulate the "Safe System" approach. While the prevention of road trauma through "safe road use, safe roads and safe vehicles" is a stated goal, the progress towards which would be assessed at the Global or First Tier level, there is no description of these elements as system "cornerstones". Further, the additional element included in current Australian conceptualisations of the Safe System, safe speeds, is specified in neither the MUARC framework nor the 2004 -2011 Queensland road safety strategy.

3.3.2 Serious injury

The MUARC framework considers severity as a part of the analyses. However, it uses the Queensland Road Crash definition of serious as "hospitalised" crashes (where the most severely injured party was reported as being taken to hospital). This is a broad classification and has been shown to be an inaccurate representation of hospitalised injuries (Watson, Watson, & Vallmuur, 2013). Discussion of the issue of defining serious injury can be found in Section 4.2.

3.3.3 Distraction and inattention

While driver distraction has long been recognised (and debated) as a contributory factor in road traffic crashes, distraction and inattention have only recently received formal recognition and greater attention in Queensland road safety campaigns. The "Fatal Four" has become the "Fatal Five" with the addition in Queensland campaigns of distraction and inattention to the other four risk factors of drink and drug driving, speeding, seatbelt non-use and fatigue. The MUARC framework, however, only makes reference to the "Fatal Four" and does not address in any detail the issues of distraction and inattention or how their contributions to crashes and the effectiveness of relevant countermeasures might be measured.

3.3.4 Drug driving

With roadside and random drug testing of drivers emerging only relatively recently as an enforcement measure, the MUARC framework contains little specific content related to drug driving. In some ways, the issue in terms of its road safety impact, the available countermeasures and the appropriate methods of evaluation shares much in common with the problem of drink driving. In theory there should therefore be no barrier to assessing the extent of drug driving and the effectiveness of countermeasures through the analysis of enforcement, crash and other data, as occurs in regard to drink driving. In practice, however, the methods for detecting drug driving are still under refinement, complicated by the need to test for multiple substances simultaneously, while there are also ongoing debates about scientific validity which do not apply to drink driving enforcement. Consequently, there is a much lower level of drug driving enforcement which, coupled with its more targeted nature, raises questions about whether or not detection figures provide reliable information about trends in actual drug driving. Arguably, reliable information about drug driving trends and prevalence requires a detection regime that is both random and comprehensive in coverage, but the case for such a program seems unclear on current cost-benefit grounds.

3.3.5 Traffic safety culture

There is no mention of "culture" or an equivalent concept in the MUARC framework. With the benefit of hindsight, this appears to be a potential weakness in the framework, although the framework need arguably only address what the strategy includes.

One element of the 2004-11 Strategy (in addition to Safe attitudes, behaviours etc; Safe road and roadsides; and Safe vehicles) is to foster "a community that values road safety as a priority" (p.29). This appears to be something of a precursor to the explicit concept of "traffic safety culture". The current Queensland strategy contains limited but clearer reference to "culture", with the statement that "we will drive a fundamental change in the culture and attitude to road safety". This refers to "cultural change among stakeholders", which is somewhat open to interpretation (may or may not include road users), rather than something like traffic safety culture as a broader philosophy. Under the heading of "the system we want by 2021", there are four objectives relating to Safe System cornerstones (or pillars) to be achieved in regard to "our culture".

3.3.6 Regional and remote classification

As mentioned above the classification of rural and urban in the existing framework is possibly too broad to capture the distinct differences in "rural" locations in Queensland. As a result, it is suggested that the Australian Bureau of Statistic's Australian Standard Geographic Classification (ASGC) Remoteness Structure be used. This system broadly classifies geographic areas based on their distance from the five nearest major population centres (National Centre for Social Applications of GIS, 2009). It is categorised into five groups (1 = Major Cities; 2 = Inner Regional; 3 = Outer Regional; 4 = Remote; 5 = Very

Remote). As shown in Table 3.1, compared to states like Victoria, New South Wales, and South Australia, Queensland has a larger proportion of the population living in Outer Regional, Remote, and Very Remote locations. The impact of programs, strategies and actions in the Road Safety Strategy are expected to be different in these locations (e.g., enforcement programs in remote and very remote areas compared to inner regional and major cities). As a result, it is important that analyses include these classifications in order to best understand how the strategy is working and areas for improvement particularly in the next action plan cycle. It should be noted, however, that for some analyses at particular levels or strata, the two remote categories may need to be collapsed together (or even possibly included with Outer Regional) for the sake of statistical power.

Table 3.1: Proportion of the population by remoteness area in each Australian state and territory

State/Territory	Major Cities	Inner Regional	Outer Regional	Remote	Very Remote
Queensland	62.0%	20.3%	14.7%	1.7%	1.3%
New South Wales	74.2%	19.3%	6.0%	0.4%	0.1%
Victoria	76.6%	19.1%	4.3%	0.1%	0.0%
South Australia	73.4%	10.9%	12.1%	2.7%	0.9%
Western Australia	76.7%	9.1%	7.5%	4.0%	2.7%
Tasmania	0.0%	65.7%	32.3%	1.5%	0.5%
ACT	99.8%	0.2%	0.0%	0.0%	0.0%
Northern Territory	0.0%	0.0%	56.6%	20.7%	22.8%

Recommendations for adaptation and/or additions to the existing framework are described in the next section.

4 RECOMMENDATIONS TO UPDATE THE FRAMEWORK IN LIGHT OF THE TARGET, VISION AND OBJECTIVES OF THE 2015 – 2021 STRATEGY

4.1 Traffic safety culture

4.1.1 Introduction

Traffic safety culture (TSC) is a relatively new concept that has its roots in organisational safety culture. As will be outlined below, TSC as a concept draws on a psychological approach to culture (rather than a broader anthropological approach) that links the beliefs, attitudes and values shared by individuals with the systems and structures within which road user behaviour takes place. Norms are a key element of the beliefs TSC focuses on, i.e. people's beliefs about what behaviours their friends, peers or society would approve of. Measurement of TSC is now a central feature of the US Government's evaluation of progress in road safety, measured through periodic surveys of public knowledge, attitudes, and behaviours relating to traffic safety. The overarching aim of this commitment is to transform what is considered to be a culture of "complacency" or "indifference" when it comes to traffic safety (Hedlund, 2007) – a complacency not seen in other facets of life, or indeed within other areas of transport such as aviation or train travel.

Evidence from the US suggests that real progress in road safety depends on more than developing and implementing specific countermeasures for specific behaviours, it requires changing the culture of indifference (Hedlund, 2007). As a result, new approaches are required to see further improvements in road safety. Recent research suggests that the application of safety culture to traffic safety may have great potential for improving road safety (Atchley, Shi & Yamamoto, 2014). This is acknowledged in the Queensland Road Safety Strategy: "to drive a fundamental change in the culture and attitude to road safety". As stated by Girasek (2012. p. 45):

> *Just as the field [of road safety] calls for scientific evidence before endorsing an engineering countermeasure, we must take an evidence-based approach to creating environments in which traffic safety is highly valued and rigorously pursued*.

This review will explore the application of safety culture research to the road safety domain, without going into detail. A more detailed analysis of the evolution of the TSC concept from the organisational safety culture concept, and the implications this has for application of TSC in practice, can be found in Edwards, Freeman, Soole and Watson (2014).

4.1.2 Defining traffic safety culture

As noted above, TSC as a concept has its roots in the "safety culture" concept used in relation to organisational culture. Much of the research on safety culture has been conducted in an organisational setting, particularly relating to high risk industries such as nuclear power and aviation, and its origin can be traced to the Chernobyl nuclear accident in 1986 in which a "poor safety culture" was identified as a contributing factor (Edwards et al, 2014). Although safety culture within transport companies was addressed in a significant (though small) body of organisational research, TSC as a wider, community-level concept was arguably first highlighted as a research priority in the United States through an AAA Foundation for Traffic Safety workshop in 2006. A compendium consisting of 22 articles was commissioned to address TSC: *Improving Traffic Safety Culture in the United States: The Journey Forward*. This formed the bulk of the existing literature on TSC up to 2006, following which the field has expanded.

In 2010, a US White Paper was dedicated to TSC as part of an effort to develop a US National Strategy on Highway Safety (Ward, Linkenbach, Keller & Otto, 2010). The paper states that (p. vii):

> *Traffic safety culture appears to be an intuitive and powerful concept with which to explain observed differences in international, regional and demographic crash risk, as well as the propensity to commit high risk behaviors. If it is possible to define and apply this concept within a relevant social psychological theory of behavioral choice, it may be possible to develop a new paradigm for traffic safety interventions.*

In the same document, Ward et al (2010, p. 4-5) provide a definition of TSC that can be presented (with minor modifications) as:

> *…perceptions people have about what behaviours are normal in their peer group and their expectations for how that group reacts to violations to these behavioural norms. In terms of traffic safety, this definition applies to behaviours that either increase crash risk (e.g. speeding) or are protective (e.g. wearing seatbelts), as well as behaviours related to acceptance or rejection of traffic safety interventions.*

This definition explicitly mentions peer groups but not the broader context of community or national TSC. To address this, Edwards et al (2014) emphasise the importance of exploring shared beliefs, attitudes, and values (cultural factors) and relevant community structures and systems (contextual factors) which influence safety-related behaviours. They provide the following definition (p.296):

> *Traffic safety culture is the assembly of underlying assumptions, beliefs, values and attitudes shared by members of a community, which interact with a community's structures and systems to influence road safety related behaviours.*

The term "community" in the definition can be interpreted in ways which encompass major stakeholder groups and the general community, at both a national and local level. Measuring TSC at the national level, state level, or within specific groups therefore involves measuring assumptions, beliefs, values and attitudes that influence road safety behaviours, within the context of legislation, enforcement, education etc.

Edwards et al (2014) also highlight an important distinction between approaches to organisational safety culture which is relevant to TSC and to its measurement. Citing Nævestad (2009), they distinguish between *interpretive* and *functionalist* approaches to safety culture. The interpretive approach emphasises to need to understand deep cultural meanings, beliefs, attitudes and values which are seen to motivate behaviour; the more dominant functionalist approach emphasises changing behaviour, or culture, through the use of organisational structures and systems. While both entail an organisational safety culture which positively influences safety by prioritising safety-related beliefs, values and attitudes (Edwards et al, 2014), the functionalist approach lends itself to the measurement of culture and culture change via standard instruments such as questionnaires, whereas the interpretive approach requires a primarily qualitative approach. In terms of the evaluation framework for the Queensland Road Safety Strategy, a functionalist approach is easier to implement.

4.1.3 The elements of TSC

As with the definition of TSC, there have been attempts to identify elements of TSC as an extension of the elements of organisational safety culture. For example, four essential elements of safety culture are proposed by Wiegmann et al (2007): *organisational commitment* to safety, the *involvement of operational supervisors* in safety-related activities (although *engagement* is a better label for their description of this element), the *formal safety system* of the organisation, and the organisation's *informal safety system*. Wiegmann et al then attempt to demonstrate how these elements can be applied to TSC by considering how they would apply at community level rather than within an organisation:

- *Organisational commitment* to safety: in the community, the government can be considered the equivalent of senior management in an organisation (city, state, and federal). Government decision makers are responsible for ensuring adequate safety resources and sufficient law enforcement personnel. They also set policy and laws regarding safety. They also make the point that the values and priorities of federal and community leaders are typically conveyed to the public through the actions they take. For example, if it takes a fatal crash to prompt action at particular intersection with a high crash rate, the public will see the commitment to safety as low.

- *Operational supervisors' involvement (engagement) in safety*: most drivers do not have formal supervisors, however some public employees are responsible for

monitoring safety and setting a positive safety example – i.e. law enforcement officers, public transport personnel, those who provide driver education and training, and those responsible for testing, evaluating, and licensing drivers. People working in law enforcement and public transport can set a positive example of safe behaviour. A positive safety culture at this community level also "might best be expressed through consistent and fair monitoring and enforcement of all safety-related behaviours" (p.121). For example, where there are gaps in enforcement, drivers think they can "beat the system" rather than being encouraged to drive safely. Those working in driver education, training, and testing/evaluating drivers also maintain a level of supervisory capacity and whose behaviour might also encourage a positive safety culture.

- *Formal safety system:* in relation to traffic safety, law enforcement is part of this system. A strong positive safety culture offers additional mechanisms for reporting hazards, i.e. people notifying authorities about traffic light power outages, taking part in research about driving behaviour, and complaining about a poorly maintained stretch of road.

- *Informal safety system:* the unwritten norms regarding safety, and possibly the greatest challenge to TSC. Changing the unwritten norms and peer culture can be difficult, especially because there are likely to be many. However, the success of safety interventions relies on it. Changing specific behaviours such as drink driving and seat belt use often involves lengthy public education and enforcement campaigns. The long term success of certain road safety campaigns indicates that changing the informal safety culture is possible.

The informal safety system has clear application to many of the statements in the Queensland Road Safety Strategy and the values, attitudes, beliefs and behaviours of communities and peer groups of various categories of road user. The other elements are highly relevant to road safety stakeholders. None of the elements are unfamiliar to road safety researchers and practitioners, however the focus on transformation of culture makes their interdependence more evident.

4.1.4 Variations in TSC and their relevance to evaluation of the Strategy

The discussion above mentions peer influences and referents, community-level TSC and the engagement and commitment of stakeholders, all of which imply that there may be many TSCs in practice, i.e. that values, attitudes, beliefs and behaviours of one stakeholder group, or peer group of road users, etc. may differ from another. Again, this is not a novel finding and is widely acknowledged in approaches to road safety campaigns, however it highlights that promoting a desired TSC entails the same processes of market segmentation and targeted campaigns that characterise best practice approaches to road safety public education. In turn, this has implications for an evaluation strategy, which can be presented as follows:

- Evaluating the prevalence of a desired TSC: from an overall road safety strategy evaluation perspective, it can be argued that the main aim is to measure the degree to which a desired TSC for Queensland has been internalised; this is a relatively simple process, since a standard list of items could be assessed (agreement with statements of desired attitudes, self-reports of desired behaviours, etc.). If this was all that was required, the evaluation instrument does not need to take into account the variety of TSCs that might exist.

- Evaluating progress among identified high risk/high interest subgroups: when the strategy has identified a high risk subgroup (e.g. young drivers) or a high interest subgroup (e.g. a stakeholder type such as road engineers), it becomes important to measure progress towards a desired TSC for these particular groups. While much of this could be achieved through structured sampling approaches using the same standard list of items as above, there may be a need for tailored items specific to issues among that subgroup. The question for an evaluation framework is whether it is feasible or necessary to incorporate this level of detail in the instruments that are a part of the framework.

The first of these objectives is a fundamental need in an evaluation framework, whereas the second objective presents problems of resourcing, complexity and change (since the groups of high risk or interest may change over time). It is suggested that a compromise solution is a standard set of items that provide evidence for the prevalence of a desired TSC, with a sampling approach that ensures that known high risk groups are sampled sufficiently to provide reliable information on their progress on the standard indicators only. This would not apply to stakeholder groups, who would need to be sampled in a different way.

4.1.5 Measurement

Given the origin of the TSC approach in organisational safety culture, the measurement approach used in the organisational setting can provide guidance on measurement of TSC. Organisational safety culture is generally measured by self-report safety climate questionnaires (note that there is debate about the distinction between “climate” and “culture”; typically “climate” is considered to be a current snapshot of a more enduring culture, but this argument is not particularly sustainable). The measurement of organisational safety culture was driven by a move towards use of more “leading indicators” rather than what was a previous reliance on “lagging indicators” such as number of fatalities and accident rates. Leading indicators are argued to be more predictive measures of safety, reducing the need to simply wait for the system to fail in order to identify weaknesses and take remedial action (Flin, Mearns, O’Connor & Bryden, 2000). The shift was driven by the awareness of human, management and organisational factors, rather than focusing on purely technical failures as the primary causes of accidents. A range of features are assessed when measuring safety culture, typically perceptions of (or attitudes toward) the current state of some facet of the organisation,

self-report behaviours, personality, and measures of individual dispositions. Common themes relate to management, safety systems, risk, work pressure, and competence. Specific items are variable and likely to be industry or even company-specific (Flin et al, 2000).

Along with the growing interest in the concept of TSC, there has been some attempt to measure the construct. Since 2008, the AAA Foundation has administered the annual Traffic Safety Culture Index. The survey measures public knowledge, attitudes, and behaviours relevant to traffic safety and tracks them over time. It includes questions about personal experience with crashes, perceptions of safety, attitudes, and behaviours. The latter category includes questions about drinking and driving, cell phone use, and texting, speeding, red light running, drowsy driving, and seatbelt use (Nævestad & Bjørnskau, 2012).

Girasek (2012) examined the constituent components of TSC and developed corresponding survey items to measure national TSC, as reflected by self-reports of the general public. Used a Delphi Technique, survey items were developed and mailed to a representative sample of 1700 US households. It measured beliefs, attitudes and self-report behaviour. Factor analysis was then performed on the 750 completed survey results. She argued for the need to look further than the level of the individual and included respondents' views on issues that went beyond driver's behaviour (e.g. traffic safety law enforcement, funding levels, road maintenance, evidence based interventions, policy making, injury surveillance). The idea was that this would gauge public support for evidence based traffic safety broadly. Moeckli & Lee (2007) stated that any approach to shifting TSC must move beyond a singular focus on the driving public, reinforcing the shared responsibility of other sectors (e.g. law enforcement, policy makers, educators, engineers) (cited in Girasek, 2012). Results of the factor analysis showed that government involvement emerged most prominently among the TSC's identified factors. This is encouraging given that public policies are associated with the most effective injury prevention strategies (e.g. speed limits and vehicle standards). This emphasis is also consistent with the "Vision Zero" approach, whereby transportation systems' designers are also accountable for safety.

The three factors that followed "increased government attention" were strict monitoring and control of alcohol-impaired drivers, disapproval of speeding, and avoidance of aggressive driving. Other factors included local engagement, desire for government and private sector accountability, more information, school involvement, teen restrictions, willingness to invest, and seatbelt use. The identified factors were in line with other TSC descriptions. In their study that explored cultural differences between rural and urban samples, Rakauskas (2009) included items that measured driving behaviours, and perceptions of policies, enforcement, engineering and educational interventions. Respondents didn't appear to think about "safe driving" as one entity – they responded differently to questions about different driving behaviours (i.e. aggressive driving, speeding/distracted/fatigued driving, and drink driving questions). While the

hypothesised model's components were categorised according to strategy (i.e. public policy, surveillance), respondents appeared to think about TSC by hazard (Girasek, 2012). This may relate to what Edwards et al (2014) highlighted; that aspects of changing TSC appear to be behaviour-specific.

In sum, international experience shows there is a need to develop a valid and reliable measure of TSC. This could then be used to identify beliefs and attitudes that feed into poor culture and at risk subgroups. In turn, this would inform future interventions and evaluation and enable TSC trends to be tracked over time. Girasek (2012) recommends that TSC surveys should include items that focus on the behaviour of policy makers and other public servants – instead of implying that drivers are the only group with the power and responsibility for reducing deaths and injuries. So, rather than simply asking individual members of the public about their knowledge, attitudes and self-report driving behaviours, the survey should extend questions to also ask about how individuals perceive various road safety stakeholder groups in order to gauge how this might influence TSC. For example, what effect different road safety stakeholders have on driver behaviour (e.g. the police), how road safety stakeholder groups interact with the public, and how they interact with each other.

From an Australian perspective, there is a relative homogeneity of Australian culture across geographical areas, and a tool for measuring TSCs in Australia therefore needs to sample those aspects of Australian culture relevant to TSC. One approach to understanding culture is that of Hofstede (1984), who identified four key "cultural dimensions" (individualism/collectivism, power distance, uncertainty avoidance, and masculinity/femininity) that have been found to explain many observed differences in work behaviours. More recently, additional dimensions such as indulgence and long-term orientation have been added. Edwards et al (2014) discuss Hofstede's model in relation to Australia. According to the model, Australia scores low in power distance and high in individualism, meaning that Australians expect a more equal relationship with those in authority, and are expected to look out for themselves. Australia also scores high on indulgence, meaning that people seek to enjoy life and wish to act and spend money in a way that pleases them. It is also considered to be a masculine society (people strive to be their best, rather than for quality of life and care for others) and low in uncertainty avoidance (people are willing to take risks and try new ideas and technologies). Taken together, it can be hypothesised that Australians may be more resistant to on-road enforcement (low power distance and high individualism), more likely to engage in sensation seeking (high indulgence), more influenced by safety messages directed at personal costs and benefits (high masculinity and individualism) and more willing to accept new types of road safety initiatives (low uncertainty avoidance) (Edwards et al, 2014). These cultural dimensions could be among those sampled in an evaluation framework. These broader dimensions are unlikely to change quickly, however their importance lies in the contextual information they contribute to the process of developing interventions that achieve changes in beliefs, attitudes and behaviours. For example, the

degree to which Australians are oriented towards sensation seeking and individualism will influence the degree to which enforcement-based approaches and deterrence-based public education will be effective, in comparison with community approaches that emphasise social responsibility. In contrast to these more stable dimensions, specific attitudes to road safety and road use behaviours can change more quickly (as longstanding experience in Queensland and elsewhere demonstrates), and therefore measurement of these specific attitudes and behaviours acts as a useful short term indicator of progress towards long term change in the more stable cultural dimensions.

Looking specifically at road safety strategy in Australia, the use of the Safe System approach means that a desired TSC would include "systems thinking", i.e. an awareness that road safety is the product of the interaction of system characteristics (road infrastructure, legislation and regulation, enforcement practices, vehicle standards, etc.) and road user behaviour, with the corollary that successfully addressing road safety requires a system response rather than isolated initiatives. This is a more challenging approach to take, as the balance between attribution of blame for crashes to system failure and acceptance of personal responsibility need to be balanced.

For example, the major contribution William Haddon sought to make to road safety in the 1960s and 1970s, through promotion of the matrix he had been applying to other areas of injury analysis, was to take away the focus on individual responsibility and emphasise system changes (Murray et al, 2014). At the time there were many unsafe vehicles on the road and unsafe road infrastructure, but blame for crashes was attributed to drivers. Haddon's argument was that driver behaviour was very difficult to change, while system changes were more achievable (Haddon, 1972). Ironically, when Hedlund (2007) synthesised the TSC literature in the 2006 AAA Foundation for Traffic Safety workshop mentioned above, one of the four culture problems he identified was the focus on safe vehicles and roads rather than on safe driving behaviour. This is associated with the notion of complacency or indifference Hedlund found was a dominant feature in the compendium commentaries on TSC. That is, the approach considered by Haddon to be a strength was, more than three decades later, considered to be a weakness.

The underlying issue in this apparent contradiction is related to whether or not people see themselves (as road users) as part of the system. If "the system" is an external set of infrastructure, regulations and practices for which people "other than me" are responsible, i.e. government agencies, legislators, vehicle manufacturers, then it is possible to "blame the system" and absolve oneself of any personal responsibility. However, the Safe System approach in Australia, and in related strategic approaches elsewhere (e.g. the United Nations Decade of Action Plan for Road Safety) incorporate road users as part of the system.

It is therefore important that measurement of TSC incorporate measurement of the degree to which road users see their own behaviour as being part of the system, i.e. a sense that they are partners in the overall achievement of a Safe System. The measurement of

this kind of "systems thinking" is not widely canvassed in the literature. A Systems Thinking Scale has been developed (Moore et al, 2010), however its use and validation across different systems areas is very limited. Some reasons for this can be seen in the initial bank of survey items considered (Figure 4.1), where the context (healthcare professionals and students) had a strong influence on the nature of the questions.

1.	I think the harder people work the better the outcomes will be.
2.	I seek everyone's view of the situation.
3.	I look beyond a specific event to determine the cause of the problem.
4.	I think understanding how the chain of events occur is crucial.
5.	I include people in my work unit to find a solution.
6.	I think outcomes are random.
7.	I think that lasting change relies on personal effort and motivation.
8.	I think recurring patterns are more important than any one specific event.
9.	I think of the problem at hand as a series of connected issues.
10.	I consider the cause and effect that is occurring in a situation.
11.	I consider the relationships among co-workers in the work unit.
12.	I think that systems are constantly changing.
13.	I propose solutions that affect the work environment, not specific individuals.
14.	I focus on my first idea because it is often the best.
15.	I keep in mind that proposed changes can affect the whole system.
16.	The main reason for success is to get the person in charge to change.
17.	I think my first impressions turn out to be very useful.
18.	I think more than one or two people are needed to have success.
19.	I keep the mission and purpose of the organization in mind.
20.	I think small changes can produce important results.
21.	I consider how multiple changes affect each other.
22.	I focus primarily on the opinions of a champion in the system.
23.	I think about how different employees might be affected by the improvement.
24.	I think personal commitment is important in creating lasting change.
25.	I try strategies that do not rely on people's memory.
26.	I recognize system problems are influenced by past events.
27.	I think that system-wide change is easy to accomplish.
28.	I consider the past history and culture of the work unit.
29.	I consider that the same action can have different effects over time, depending on the state of the system.
30.	I think uncertainty and surprise are involved.

Figure 2: Systems Thinking Scale initial item bank (Moore et al, 2010, Fig. 3, p. 6)

It is therefore important that a suitable scale, appropriate to the Safe System approach, be developed. It would include some items similar to those in the Systems Thinking Scale, but would need to be tailored to the characteristics of the Safe System. It would also need to be relevant to stakeholder groups, e.g. to sample the "Safe Systems thinking" of engineers, police and regulators as well as road users in general.

Figure 2 presents a roadmap as an approach to the measurement and use of TSC, based on a proposal by Girasek (2012). For the evaluation framework for the Queensland Road Safety Strategy, the key output of the approach is the development of the instrument, which can then be used for periodic monitoring and evaluation.

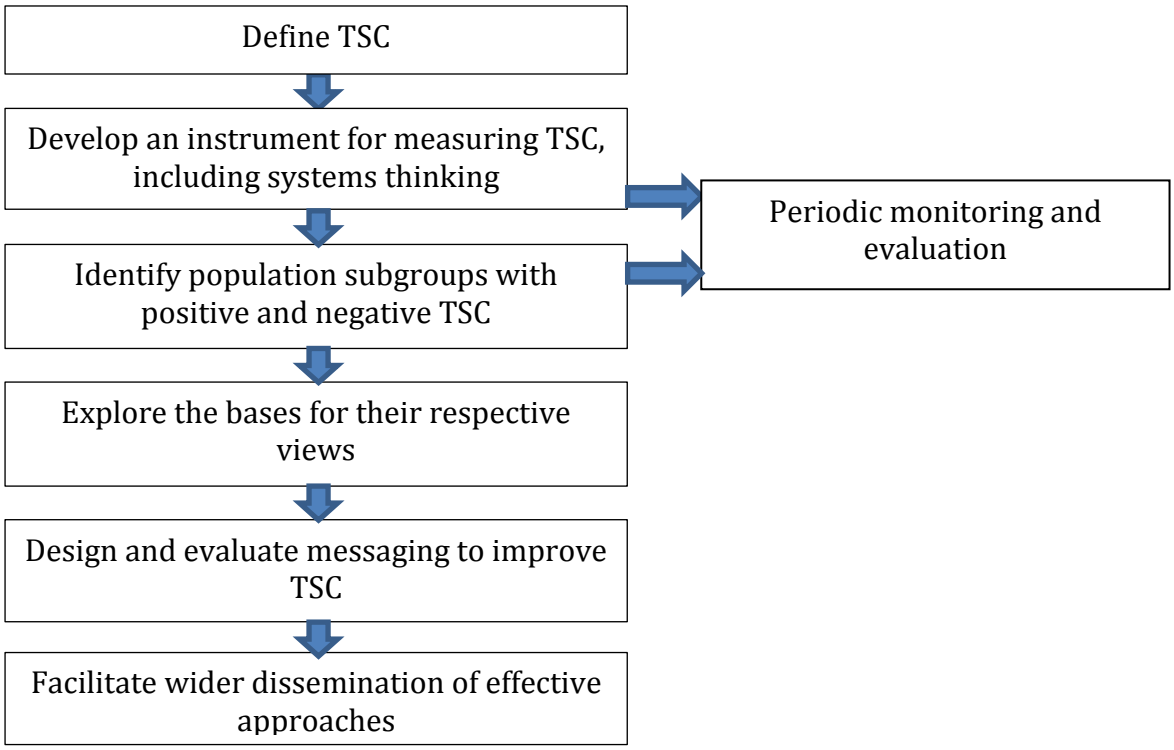

Figure 3: Advancing TSC: a research road map (Modified from Girasek, 2012)

4.1.6 Summary

TSC may hold great promise for improving road safety. Conceptually, organisational safety culture provides a useful paradigm for consideration of TSC – at both a national and community/peer group level, and in a more direct way for road safety stakeholder organisations to look at their own internal safety cultures. As was the case for organisational safety, looking at TSC can be seen as more of a leading indicator of safety, rather than simply focusing on fatalities and injuries. There is no single strategy for changing TSC. It requires many actions and strategies at all levels, some broad, and some specific. This involves rethinking values, beliefs and norms relating to appropriate behaviour on the road taking into account that many of them lie outside the domain of TMR or even government responsibility, and may be reinforced or perpetuated by structural factors related to economic imperatives, infrastructure constraints and political priorities. Improving TSC is a way to modify peoples' existing attitudes, beliefs and knowledge that may be feeding into poor culture and unsafe behaviour. By continuing to improve the culture, these attitudes, beliefs and knowledge will continue to change. As Edwards et al (2014) argue it may never be possible to create an ideal culture; however, investigating an existing culture can enable the identification of effective safety strategies.

Measurement of TSC should go beyond simply measuring public knowledge in terms of their individual attitudes, beliefs and self-report driver behaviours; it should also involve measurement of Safe Systems thinking and their perceptions of relevant road safety stakeholder groups in so far as they influence road user safety.

4.2 Serious injury and under-reporting

4.2.1 Introduction

Over the years there have been significant reductions in fatalities in Australia (The Parliament of Victoria Road Safety Committee, 2014), as there has been in many other highly motorised countries (International Traffic Safety Data and Analysis Group (IRTAD), 2011). However, there has been less of a reduction (and in some cases an increase) in the number of serious non-fatal road crash injuries in many of these jurisdictions, including Australia. This in combination with the substantial burden of serious non-fatal road crash injuries has meant that nationally and internationally, the focus in road safety has shifted towards a greater understanding of road crash serious injuries in addition to fatalities (International Traffic Safety Data and Analysis Group (IRTAD), 2011; The Parliament of Victoria Road Safety Committee, 2014). Police reported crash data are the primary source of crash information in most jurisdictions (International Traffic Safety Data and Analysis Group (IRTAD), 2011). Unfortunately, however, the definition of serious injury within police-reported data is not consistent across jurisdictions and may not be accurately operationalised, which could lead to misleading estimates of the impact and cost of crashes.

The definition of a fatality is relatively consistent across countries since it is based on that of the Organisation for Economic Co-operation and Development (OECD) whereby a traffic death is one that occurs within 30 days of a road crash (OECD, 2016). In terms of other severity levels, particularly in relation to serious injury, the definitions are much more variable. Many of the countries in the OECD define a serious injury as a person who is admitted to hospital for 24 hours or more as a result of a road crash (World Health Organization, 2010). However, this definition generally relies on the police identifying injured persons based on whether they were transported to hospital or not. Given the reported lack of liaising between police and hospitals on the length of admission, a serious injury category using this definition could range from cuts and bruises to severe head injuries (Ward et al., 2010). Another study by the authors (Watson, Watson, & Vallmuur, 2013), suggest that the "hospitalisation" severity category used by police does not reflect true hospitalisations in all cases, at least within the state of Queensland. Further, it highlighted the wide variety of severity levels within hospitalised cases that are not captured by the current police-reported definitions.

As a result of this broad, and likely inconsistent, serious injury classification, more objective and precise measures of severity have been proposed (International Traffic Safety Data and Analysis Group (IRTAD), 2011) which rely on either police assigning a nature of injury code or on the use of hospital discharge diagnoses (e.g., Abbreviated Injury Scale, ICISS). The Abbreviated Injury Scale (AIS) is a body-region based coding system developed by the Association for the Advancement of Automotive Medicine (Association for the Advancement of Automotive Medicine, 2008). A single injury is classified on a scale from 1-6 (1 = minor; 2 = moderate; 3 = serious; 4 = severe; 5 = critical;

and 6 = maximum). Another example of a more precise measure of severity is the International Classification of Diseases–based Injury Severity Score (ICISS) (Osler, Rutledge, Deis, & Bedrick, 1996). ICISS involves using ICD diagnoses to calculate threat-to-life associated with an injury. Survival Risk Ratios (SRR), which are the proportion of cases with that diagnosis code which did not die, are calculated for each ICD diagnosis code. Cases are then assigned an ICISS, which is the multiplication of SRRs of all their diagnoses. It should be noted that there is some debate surrounding the most appropriate injury severity classifications, however these two measures are widely accepted and often used in injury research as reasonably reliable measures of the probability of death (Langley & Cryer, 2012; Stephenson, Langley, Henley, & Harrison, 2003).

It could be suggested, however, that even if more detailed information was collected in order to assign these more objective and/or precise measures, the police are not necessarily in the best position to collect this information. Police do not have the training or expertise to record information on the nature of an injury, or injuries, with the required level of accuracy. Also, even if they were trained to assess this, classifying injury at the scene of a crash could be problematic, as not all injuries are apparent at the scene and the police have many competing priorities in these situations (e.g., traffic control). Also, it is argued that the consistency of the recorded information from case to case could be questionable if collected by the police (Amoros, Martin, Chiron, & Laumon, 2007; Chapman & Rosman, 2008; Farmer, 2003; McDonald, Davie, & Langley, 2009; Ward et al., 2010). The World Health Organisation (2010) suggests some possible strategies for addressing the issue of serious road crash injuries, including data linkage between police and hospital databases either routinely or periodically to check the accuracy of the police data, and/or the following up of cases by police (or reported by the hospital) to determine the length of the hospital stay.

4.2.2 Hospital data

The hospital data referred to in this section is the Queensland Hospital Admitted Patients Data Collection (QHAPDC). QHAPDC contains data on all patients discharged, statistically separated, died, or transferred from a Queensland hospital permitted to admit patients (including public hospitals, licensed private hospitals, and day surgery units). According to the QHAPDC manual, generally "a patient can be admitted if one or more of the following apply:

- The patient's condition requires clinical management and/or facilities are not available in their usual residential environment
- The patient requires observation in order to be assessed or diagnosed
- The patient requires at least daily assessment of their medication needs
- The patient requires a procedure(s) that cannot be performed in a stand-alone facility, such as a doctor's room, without specialised support facilities and/or expertise being available
- There is a legal requirement for admission (e.g. under child protection legislation)

- The patient is aged nine days or less." (Queensland Health, 2012, p. 32)

Under the National Healthcare Agreement (NHA) between the Australian government and the State of Queensland, hospitals permitted to admit patients must provide information about admissions to QHAPDC. These data are used for a number of purposes including monitoring funding arrangements, requesting additional funding, epidemiological study (morbidity and mortality), education of students of medicine, nursing and allied health. QHAPDC is housed on a secure server within the Health Statistics Centre (HSC), under the governance of Queensland Health.

Data is collected in each of the facilities included in the collection. Data is collected in two ways depending on the hospital, either the Hospital Based Corporate Information System (HBCIS) or a paper based system (Identification and Diagnosis Sheets and Patient Activity Form). HBCIS data are extracted and mapped to the Data Collections Unit requirements, the translation of which is outlined in the QHAPDC manual. Data is collected monthly in unit record form. If forms are used, they are sent to the Area Health Service to be converted into approved electronic format and then forwarded to the Data Collections Unit (HSC). HBCIS data is sent directly to the Data Collections Unit (HSC).

Different elements of the data are collected by different staff. Admitting staff collect the following:

- Unique Record ID
- Facility name and number
- Queensland Ambulance number (eARF number)
- Admission date
- Admission time
- Date of birth
- Sex
- Patient family and given names
- Patient address
- Compensable status
- Country of birth
- Indigenous status
- Nature of injury.

Discharge staff complete the following:

- Separation date
- Separation time
- Mode of separation.

Medical practitioners complete the following:

- Principal diagnosis
- External cause
- Place of occurrence.

Data is coded at the facility as well as at the Data Collections Unit. At the facility, trained data coders code clinical details using the current version of the ICD-10-AM. At HSC, data may be coded in different ways for the release of data to external parties (e.g., collapsing categories to prevent possible identification, assigning ARIA+ classifications).

The HSC checks for errors including valid values, logical consistency and historical consistency. Validation reports are produced for the hospital, in which the hospital will make corrections and resubmit to HSC. A record of these procedures conforms to the Australian Classification of Health Interventions. Data can be modified by the hospital up to September of the year after the financial year to which the data relates. It should also be noted, that when requests for data are fulfilled by HSC, further coding or re-coding may occur to fit with the need of the requesting party or to comply with legislation.

Data is required to be sent to HSC within 35 days after the month of separation. Data is subject to validation checks by HSC and reports are sent back to hospitals for correction. This process of submission, validation, correction and re-submission can take up to 8 weeks. The data is not considered final, and therefore able to be released to external parties, for several months after the end of the financial year in which the episode occurred.

Selection of cases

If hospital data is to be used for road safety monitoring and evaluation, it is important that the selection of hospital cases best matches that of a definition of a road crash. In order to do this, the first step in selection is to select land transport related external cause codes (based on the International Classification of Diseases – 10th Edition – Australian Modification). These codes are as follows:

- Pedestrian injured (V00-V09)
- Pedal cyclist injured (V10-V19)
- Motorcycle rider injured (V20-V29)
- Occupant of three-wheeled motor vehicle injured (V30-V39)
- Car occupant injured (V40-V49)
- Occupant of pick-up truck or van injured (V50-V59)
- Occupant of heavy transport vehicle injured (V60-V69)
- Bus occupant injured (V70-V79)
- Other land transport (V80-V89).

Then it would be necessary to select only those that are defined as "traffic". This is done by using the fourth character in the ICD-10-AM external cause code. Other exclusions also need to be made due to cases not fitting the definition of a road-crash. Specifically, when the injury resulted from a pedestrian colliding with a pedestrian conveyance (V00) or a railway train (V05) it should also not be included. Also, all transfers, as identified by separation mode need to be excluded to partly eliminate multiple counts of cases (Berry, Harrison, & Bureau, 2008).

It should be noted that the primary purpose of the hospital data collection is not for the surveillance of road crash injuries, As a result, while these criteria will minimise the chance of incorrectly identifying a road crash case, it is still possible. Conversely, there is also a chance that some cases may be missed. The coding of the data relies on the information provided on the case/medical files collected throughout a patient's episode of care. It is possible that some information required for accurate coding of road crash injuries is not always available and thus the coding may be unspecified as to the nature of an injury. This is particularly the case with the coding of "traffic", which, while analogous to the on-road definition in the road crash database, has defaults when the location of the injury occurrence is unknown (National Centre for Classification in Health, 2004):

> *A vehicle accident is assumed to have occurred on the public highway unless another place is specified, except in the case of accidents involving only off-road motor vehicles, which are classified as non-traffic accidents unless the contrary is stated*.

Access

Aggregate hospitalisation data for the purposes of calculating rates and trends over time is available upon request. For these requests there is no need for ethics approval. This aggregate data could be made available per month, by age group, road user type, and remoteness classification.

If more detailed information (unit record information) is required (e.g., for mapping ICD-10AM codes to serious injury indicators) then ethical and Public Health Act approval would be required. There could be a possibility of a standing arrangement between Queensland Health and Transport and Main Roads via a Memorandum of Understanding or similar contract being entered into for the provision of data. However, this would still be de-identified data which would not allow the linking to the road crash data.

Measuring serious injury using hospital data

Resource-use based

Resource-use based measures are counts of people that make use of hospital/health services for their injury. This may be the number of people admitted to hospital, or may be more specific to the number of people admitted to hospital for 24 hours or more (an

international definition) or even longer (e.g., 14 days). It may also include the use of other health services including rehabilitation.

Advantages: easy to calculate, relatively easy to obtain, good measure of burden and elements of cost.

Disadvantages: influenced by changes in admission policy, bias in admission policy (certain individuals such as children are more likely to be admitted even for possibly non-serious injuries), may mask changes due to the broadness of the measure, potentially unreliable as a trend estimate.

Adding a length of stay element to the admission to hospital (by setting a threshold of say 24 hours or more or 14 days or more) has the potential to reduce the variability in admission policies. However, it is still not without its issues. Specifically, an initial admission for a serious road crash injury may last less than 14 days but be followed up with a long rehabilitation stay (which may not be reflected in the admission status). Also, less serious road crash injuries may exceed 14 days but only due to complications or co-morbidities unrelated to the road crash injury.

Threat to life

These measures are the probability of a person dying as a result of their injury. One advantage of these types of measures is that they are likely to be less influenced by admission policy and changes to those policies. The two most commonly discussed options for threat to life measures for road crash injuries are the Abbreviated Injury Scale (AIS) and the International Classification of Disease-based Injury Severity Score (ICISS). These two measures including their advantages and disadvantages are described below.

International Classification of Disease-based Injury Severity Score (ICISS)

In order to understand this measure, it is important to understand the ICD-10AM coding. This coding system is an international standard classification of diagnoses and external causes. The ICD is the basis for the calculation of Survival Risk Ratios which are estimates of the survival for a particular injury diagnosis. It estimates the probability of survival from 0 (no chance of survival) to 1 (100% chance of survival). The SRRs are calculated for each injury by dividing the number of patients who survive with that injury by the number of patients who sustained that injury. It is generally considered that an SRR is serious if it is equal to or less than 0.941 (Cryer & Langley, 2006). If more than one injury is known an ICISS can be calculated by multiplying the SRRs for each injury.

Advantages: Ability to map to ICD-10AM already in data, accounts for multiple injuries, can be mapped retrospectively, and comparability to other states as it is being considered (or is already in use) in Victoria, WA, NSW and the Northern Territory.

Disadvantages: limited estimation of disability, may need to be updated.

Abbreviated Injury Scale (AIS)

The AIS is an internationally agreed definition that describes the severity of injury for each of the nine regions for the body and was developed by the Association for the Advancement of Automotive Medicine (AAAM). The scores assigned to different injuries are based on medical consensus and requires detailed medical information about the injury. The AIS has two components: (1) the injury descriptor and (2) the severity score. Based on the AIS, an injury is rated from 1-6 (1 = minor; 2 = moderate; 3 = serious; 4 = severe; 5 = critical; and 6 = maximum). Where there is not enough information to assign a value, a code of 9 (not specified) is used. Generally, a score of 3 or more would be considered serious.

Advantages: widely used, familiarity with the AIS amongst trauma specialists, and consistent with the International Road Traffic and Accident Database (IRTAD) definition for "serious injury".

Disadvantages: limited estimation of disability, only valid if done by specialist coders, mapping software is limited, and expensive to collect.

Burden of disease measures

These measures combine information about the burden of an injury from mortality and morbidity by incorporating the effects on both the quantity and quality of life into a single measure. These measures provide information on the long-term impact of road crashes both on individuals and the community, which is arguably lacking in the other measures already discussed. For example, an injury to a limb would generally be considered non-serious using a threat to life measures. However, these injuries can still be quite debilitating and require long-term care.

The two most common burden of disease measures are Disability-Adjusted Life Year (DALYs) and Quality-Adjusted Life Years (QALYs). These two measures are relatively similar so the focus will be on DALYs in this next section mainly due to the DALYs being the chosen measure for the World Health Organisation.

DALYs are considered to be the number of lost years of "healthy life". DALYS combine the Years of Life Lost (YLL) and the Years Lost due to Disability (YLD). The YLL is the number of deaths multiplied by the standard life expectancy at age of death. The YLD is the number of people injured multiplied by a disability weighting multiplied by the average duration of the case until remission or death. Disability weights are measured on a scale from 0 (perfect health) to 1 (equivalent to death). These weights are produced by expert panels who draw on their own experience and surveys with samples from the population. In these surveys participants are asked to rate different health states that typically result from a particular type of injury.

Advantages: is internationally recognised as an appropriate burden of injury measure, takes into account disability burden

Disadvantages: lack of jurisdictional specific disability weights, only relates to a single injury, based on surveys of participants who have often not experienced the health states they are asked to value which may underestimate their burden, biased towards younger people as the years life lost will always be greater than older people with the same injury.

While the value of DALYs is clear, it is generally acknowledged that further research is required to apply them to an Australian road crash injury specific context.

4.2.3 Under-reporting

Another issue relating to the accuracy of police-reported data is the under-reporting of road crash injuries to police which has been shown to be biased against certain road user groups. Previous work has found that a significant proportion of road crash injuries, that required hospitalisation, are not reported to police including many cyclists (up to 80%) and motorcyclists (up to 65%) (Alsop & Langley, 2001; Boufous, Finch, Hayen, & Williamson, 2008; Watson, Watson, & Vallmuur, 2015). These vulnerable road user groups are increasing in number and are therefore likely to be under-emphasised in current road safety strategies and initiatives. It is important therefore to give some consideration to using hospital data (including linked data) to determine the level and nature of under-reporting, monitor this over time and perhaps apply adjustments to the police data to account for it. It should be noted that any expectation that non-fatal serious injury data will be as accurately recorded as fatal data is unrealistic, but quality estimates can be derived with a certain level of confidence if further work can quantify the under-reporting more precisely.

4.2.4 Data linkage

Data linkage involves the bringing together of two or more different data sources that relate to the same individual or event (National Collaborative Research Infrastructure Strategy, 2008). In principle, any datasets that contain information about individuals has the potential to be linked. Data linkage is used for a variety of reasons including data quality improvements and gaining information that a single data source cannot provide.

All data linkage centres in Australia (including Queensland) and many around the world, apply what is known as the "separation principle", by only using identifying information required for linkage without any content or clinical data. The data linkers' task is to establish links using this identifying information and assign a linkage key to each match. This linkage key is then sent to the custodians for them to extract the relevant content data and provide these data with the linkage key to the researcher. Using this separation principle approach means that those performing the linkage will be unaware of the circumstances by which any individual is included in the data collection or any details relating to these circumstances. Also, the researcher will only have the data required for

analysis without any identifying information. No entity, except the data custodian, ever has access to both the personal data and the content data. This approach is often used to preserve the privacy of the individual as well as to allow data custodians to maintain control over the data collections within their governance and is considered best-practice in Australia (Boyd et al., 2012).

Data linkage can provide the "best of both worlds" from data collections such as characteristics and circumstances of crashes from police and detailed injury information form the hospitals. This information is valuable to decision making and monitoring in road safety, however it should be noted that this would only be the case for those that link. As mentioned above, there will be some crash injuries (particularly cyclists and motorcyclists) that will not have circumstance information because they are not reported to police. Data linkage can still be a valuable tool for those that do not link, by providing information about how adjustments can be made to current reporting.

There are some complexities related to linking of health and road crash data surrounding memorandums of understanding. Also, at this stage, data linkage needs to be conducted within the Queensland Health Data Linkage Unit and would only be done on a project basis involving an external researcher (i.e., not from Queensland Health or Transport and Main Roads) to be consistent with the separation principle. This sort of project based approach could be conducted in the current environment to examine data linkage over time. The test case for linkage feasibility and technical arrangements has already be tested as part of a PhD project within CARRS-Q (Watson, 2015). Agreements may be put in place in the future and there is currently an Inter-Agency Committee (Including TMR, QPS, Queensland Health, and MAIC) that are discussing the various options. The recommendations and/or actions from this group should be monitored to inform progress in this area. In addition, an Austroads funded project testing national data linkage for road crashes is currently underway which also should be monitored to inform future direction in this area.

4.2.5 Summary

By linking hospital data with police data, improvements to both the reporting and the classification of serious injury can be achieved. Data linkage between hospital and police data has been acknowledged by relevant agencies as being an integral part of reporting serious road crash injuries. It should be noted that it is unlikely that non-fatal injury data will ever be as accurate and reliable as fatal data for reasons including:

- misreporting of road crash cases in hospitals
- lack of circumstance information for those cases that do not link with a police record
- changing admission policies
- potentially inconsistent application of serious injury definitions.

It should be noted that the specific targets relating to reductions in serious injuries in the Road Safety Strategy are based on a baseline measure that is defined as the police categorised "hospitalised" injury. Also, this baseline (and the proposed targets) are described as numbers of serious injuries rather than proportional decreases. As a result, changing the definition of serious injury within the life of the strategy may be problematic as it may "artificially" inflate the baseline, making targets unachievable or conversely produce favourable results that are not based on any real change except in definition. As a result, it is suggested that any other definitions of serious injury proposed in this revised framework should be used in conjunction (in parallel) with the existing measure rather than as a replacement.

4.3 Recommendations

In light of the review of literature, the current Road Safety Strategy, and the previous evaluation framework the following are recommended:

1. Measures of traffic safety culture and systems thinking be included in the framework

2. Hospitalisations (based on hospital data) be included (in addition to the current definition) as a measure of serious injury in trend modelling and evaluation

3. Cyclists be included as a strata (grouping) for analysis in the Second and Third Tier

4. Rural and remote classification for stratification in the Second and Third Tier be more refined and match the Australian Standard Geographic Classification (ASGC) Remoteness Structure

5. Distraction as a contributor to road crashes should have more emphasis in the Second and Third Tier

6. Improvements to the measures of distraction and its role in crashes be investigated

7. Improvements to data to better distinguish between drug and drink driving and to better quantify their relative involvement in crashes be investigated

8. Developments in traffic safety culture theory and measurements be monitored

9. Opportunities for data linkage and other data improvement strategies be monitored.

5 UPDATED EVALUATION FRAMEWORK FOR THE 2015–2021 STRATEGY

The elements of the existing framework developed by MUARC are in text boxes with additions or modifications in body of this section. As with the existing framework the proposed update to the framework will be guided by GOSPA.

5.1 The Global Assessment or Top Tier Model

The Global Assessment Model relates to the broad goals and objectives of the 2015-2021 Queensland Road Safety Strategy. Specifically, the objectives of the current Road Safety Strategy are to:

- reduce fatalities from 303 (average 2008-2010) to 200 or fewer by 2020
- reduce hospitalised casualties from 6,670 (average 2008-2010) to 4,669 or fewer by 2020.

The proposed MUARC evaluation framework described the process for evaluating the Top Tier as follows:

> At the highest level of evaluation, the performance of the strategy can be measured using an intervention type forecasting model that measures change in aggregate in road trauma levels for the situation where the road safety strategy is implemented as a package compared to a situation modelled in which the strategy is not implemented at all.
>
> The global assessment model evaluates the overall effects on crashes of the major initiatives in the strategy (Newstead et al, 1998; Newstead et al, 1995). Such a model will include variations in major socio-economic factors across time and regions of the State. Hence these influences, which are potentially contaminating influences on the apparent effects of the initiatives, will be removed. The effects of initiatives addressing relatively small target groups or relatively short time periods can also be accounted for in this global-level model, allowing their potential evaluation for the first time in some cases.
>
> This global assessment model aims to measure the effect on road trauma of the Queensland Road Strategy and associated Action Plan overall. Road trauma will be quantified in terms of fatalities, serious injuries, fatality rates and serious injury rates.
>
> The global assessment model can be formulated in two different ways depending on the time frame in which the evaluation model is being formulated relative to the implementation of the strategy. In the case of formulating the evaluation model at the commencement of the strategy, a time series model (based on state space modelling techniques) is estimated that models road trauma levels each month before the introduction of the Queensland Road Safety Strategy (i.e. Pre 2004) and forecasts the levels of road trauma that would have been expected

(together with confidence limits on the estimates) to have occurred after 2004 had that strategy not been in place (based on the past trends). Against the forecasts from the resulting model can be plotted the actual road trauma trends in Queensland that occurred after the 2004-2011 Queensland Road Safety Strategy was introduced. Plotting of the actual trends against those forecast in the absence of the current strategy can be easily achieved by Queensland Transport staff without statistical training as a means of monitoring overall strategy performance over time. Figure 4.4 depicts this graphically.

The time-series model that will investigate road trauma trends pre and post the Queensland Road Safety strategy against a forecast trend post strategy in the absence of the strategy are similar in philosophy to the control-chart methodology used by a number of agencies in the past, including Queensland Transport, to monitor road safety strategy performance. The advantage of the proposed methodology for the global assessment model here is that it employs much more sophisticated and robust statistical methodology yet is still amenable to use by those without statistical training once established.

The global assessment model can also be used at various time periods after strategy implementation to formally evaluate the overall performance of the strategy to that point in time based on the key outcome measures specified in the Strategy's goals and outcomes. Here the pre and post implementation data to the time available are modelled using the state space techniques with an intervention term being included at the time of strategy implementation. The intervention term parameter then represents the effect of the strategy on the outcome measure which can be tested formally for statistical significance. The intervention term can be modified accordingly to reflect increasing effects of the broad strategy over time which might be expected if components of the strategy are introduced in a staggered manner over time or take some time to become fully effective. Application of the global assessment model in this manner would require high level trained statistical expertise.

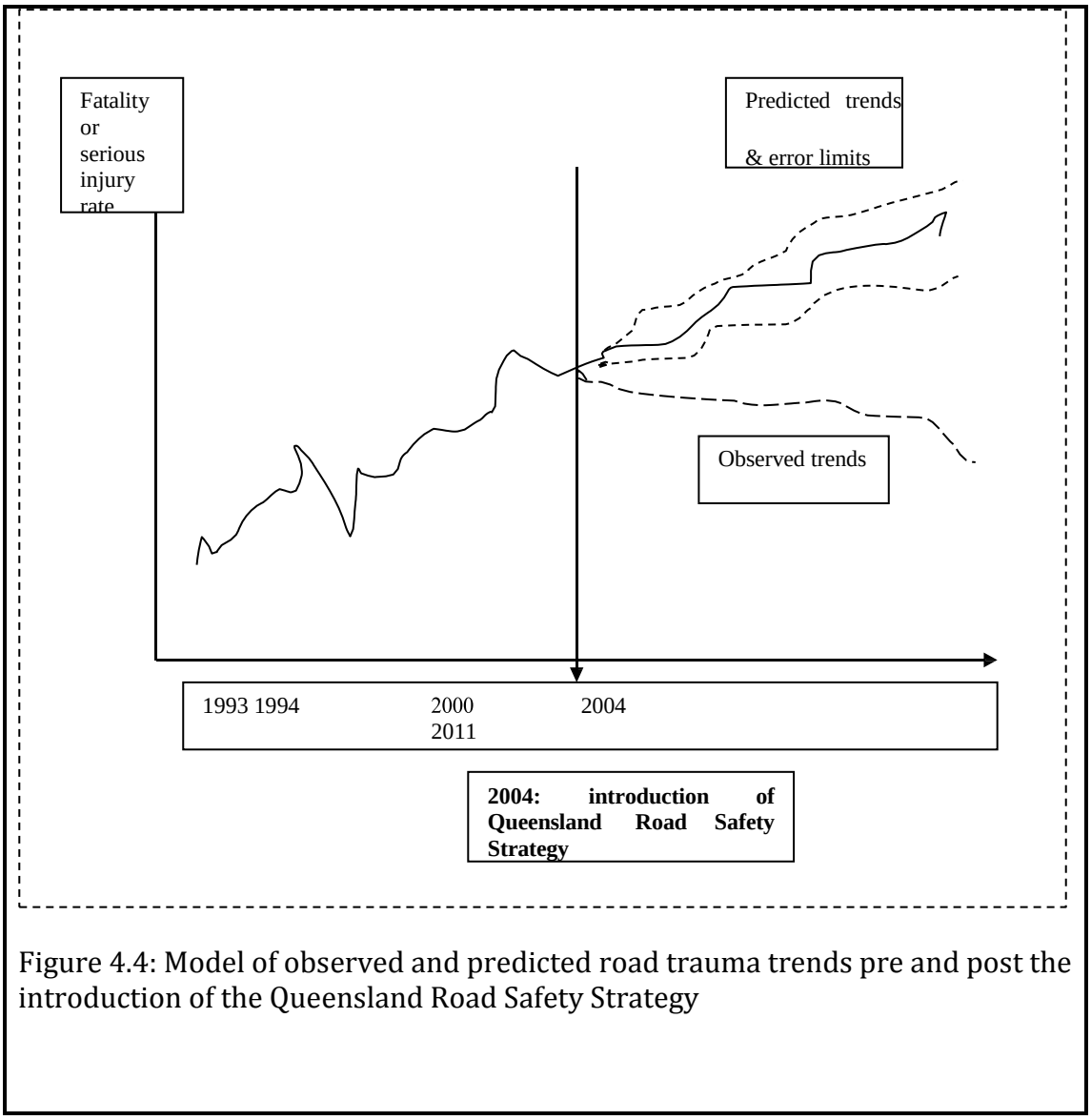

Figure 4.4: Model of observed and predicted road trauma trends pre and post the introduction of the Queensland Road Safety Strategy

It is proposed that the first tier assessment remain in its entirety. However, in addition, traffic hospitalisations in admitted patients' hospital data should be included as a supplemental measure. While ideally, linked data should also be utilised, at this stage these data are not readily available. If during the life of the strategy these data are available there should be reconsideration for their inclusion. Specifically, using the hospital data, the following counts of serious injury should be included in addition to the Queensland Road Crash hospitalisations:

- All traffic-related injuries admitted to Queensland hospitals (as defined in Section 4.2.2) for 24 hours or more
- All traffic-related injuries admitted to Queensland hospitals with an ICISS below 0.941 (based on mapping to ICD-10-AM codes).

5.2 The Second Tier Model

The Second Tier relates to the strategies and programs of the current Queensland Road Safety Strategy and the first Action Plan. As described in the text box below, the existing framework included the stratification of the First Tier analysis by broad road user groups and issues.

> Based on the 2004-2011 Queensland Road Safety Strategy and the 2004-2005 Queensland Road Safety Action Plan the broad key road user groups and behaviours/situations that need to be targeted in the second-tiered model include:
>
> - Alcohol and drug driving (modelled during high alcohol times of the week)
> - Speed-related crashes and speeding drivers
> - Fatigued drivers
> - Young adult inexperienced drivers
> - Older drivers
> - Fatal and serious crashes in rural Queensland
> - Pedestrians, including intoxicated pedestrians
> - Unrestrained vehicle occupants
> - Unlicensed drivers and riders
> - Motorcycle riders
> - Indigenous road users
> - Roads with poor crash records
> - Fleet-related crashes
>
> For each of these strata, a specific analysis model equivalent in structure to that defined by the global assessment model above would be estimated for each of the key outcomes being measured. Like the global assessment models, the second tier models can be formulated at time of implementation to forecast road trauma outcomes in each stratum of interest had the strategy implementation not taken place. Actual post implementation road trauma trends are then compared to those forecast to assess strategy effectiveness. Intervention models can also be estimated at time points after strategy implementation to formally assess the statistical significance of outcome changes related to the strategy for each strata defined above. Expertise required for each approach is the same as for the global assessment model.

In the adapted framework it is proposed that these analyses remain in their entirety, with one exception. It is proposed that the revised framework include the ASGC remoteness index (Major Cities, Inner Regional, Outer Regional, Remote, and Very Remote) as the basis for "rural roads" as defined in the existing framework as opposed to an urban/rural split. Further stratification on "rurality" is considered important in the Queensland context. Although, in some cases, the two remote categories could be combined (and even potentially combined with Outer Regional) in cases with lower numbers and statistical

power may be an issue. In addition it is proposed that the strata include cyclist injuries and distraction/inattention crashes.

Hospital data should also be used in this Second Tier analysis. However, it should be noted that not all strata will be able to be analysed using hospital data as information relating to issues such as speed, alcohol involvement, time and location of crash, and fatigue would not be available. For these strata, linked data would be required.

It would be possible however to analyse for the following strata using hospital data:

- Young adult drivers (age)
- Older drivers (age)
- Pedestrians (ICD external cause code)
- Motorcycle riders (ICD external cause code)
- Indigenous road users (indigenous status)
- Fleet-related crashes (work-related ICD activity code/compensable status)
- Cyclists (ICD external cause code)

Another element of the Second Tier that the adapted framework should include is the measures of cultural change. This would not be achievable using crash data and requires the use of pre-existing and proposed self-report measures and data. The exact details of the measures themselves are described in Section 6, but the methods for collecting these data are described below.

There are two broad groups that will need to be surveyed in order to measure the performance of the strategy on cultural change: 1) key stakeholders and 2) the broad community. The identified key stakeholders may vary, but as a starting point should include the key implementers:

- Transport and Main Roads (organisational areas/staff delivering road safety related policy, programs, public education and infrastructure)
- Queensland Police Service (organisational areas/staff delivering road safety and traffic related policy, enforcement and community engagement)
- Local government (organisational areas/staff delivering road safety related programs, infrastructure and community engagement)

Other stakeholders that should be considered include RACQ, trade unions, road user representative groups and other government departments/agencies; however such organisations are likely to have only small (or non-representative) numbers of people undertaking road safety related work. In such cases a qualitative approach could be more productive, for example a desktop review of their policy and program documents to assess the degree to which they reflect desired TSC.

A survey would need to be developed (see Section 6) and provided to representatives of these key stakeholder groups at regular intervals (e.g., yearly) throughout the life of the

strategy. Baseline data collection would not be possible (as the Strategy has already commenced), but regular surveys should allow for the detection of improvements over time. In addition, a survey "run" could be conducted before the implementation of the next Action Plan which could measure its impact over the following 2 years.

Representatives from the stakeholders should not solely be those whose position or duties relate directly to road safety, they should also include those who deal with:

- road and road-related infrastructure
- passenger transport
- customer service/contact
- data collection/incident investigation
- contact with injured road users
- active transport.

It is proposed that the current WAVE survey conducted on behalf of TMR should be the primary source of data from the broader community. The WAVE survey is annual online survey of Queensland motorists that focusses on road safety attitudes and behaviours. There are already measures within this survey that relate to concepts of traffic safety culture that can be measured for changes over time. However, there is also opportunity to include extra measures to tap into elements not currently covered. The details of pre-existing and proposed measures are described in Section 6. It should be noted that the addition of measures should be tempered against the integrity of the current survey and should not make the survey prohibitively long. Also, any additional measure will not have a baseline.

Other surveys, including those conducted by other organisations (e.g., RACQ, CARRS-Q) and the National Community Attitudes to Road Safety Survey (BITRE) should also be monitored. There also could be possibilities for dedicated traffic safety culture surveys to be developed and conducted following further research and developments in the area.

5.3 The Third Tier Model

The third tier model aims to deal with individual program elements and actions of the Queensland Road Safety Strategy and associated Action Plans.

The existing framework describes the following:

> The third tier modelling strategy is an extension of the tier 2 model in that it will typically target the same strata defined in tier 2. However, instead of modelling historic trends through general level, slope and seasonal terms, the model will include specific measures of road safety program effort under different activity areas as model covariates. In this way, the model makes estimates of the effects of individual initiatives (where there is sufficient data for the estimate to be reliable) by establishing the relationship between measurable road safety program effort

and the key strategy outcome measures and relating the real variation in program effort to the reduction in road trauma observed. Results from the tier 3 modelling process will give specific estimates of the relative contributions of each of the major program elements in the road safety strategy to achieving the measured outcomes.

Poisson or negative binomial regression models will be used in this third tier modelling approach. These will be fitted to the outcome data series at some point after program implementation when sufficient post strategy and program element implementation experience has been accumulated to allow for successful modelling outcomes. The tier 3 modelling process must be carried out by someone with high level statistical expertise.

There will be some types of initiatives whose impact on crashes cannot be assessed using the tier 3 modelling approach due to relatively small target group and/or duration of operation or because program element input cannot be measured in a meaningful way. In some instances these program will be represented as local interventions. Where this is not possible, the aggregate effects of such programs will be assessed through comparing the third tier modelling outcomes with the aggregated effects accounted for in the global assessment model.

Types of road safety activities that can be included in the tier 3 evaluation models include:

Speed camera activity:

Six measures of speed camera activity that have been found to be key predictors of crash outcomes in the full formal evaluation of the Queensland speed camera program (Newstead and Cameron, 2003) and could be used as inputs into the tier 3 models. They are:

- Total number of speed camera operation hours per month by police region
- Number of active sites available for use by police region
- Hours of operation per active camera site available for use by police region (derived from the above measures)
- Percentage of sites visited as expected according to randomised speed camera operations schedule by police region
- Monthly rate of increase in active camera sites by police region
- Monthly rate of increase in speed camera operation hours by police region.

On-road (non-speed camera) Police Enforcement

- On road speed enforcement effort by Moving Mode Radar (MR) and Laser Speed Detection (LIDAR) - operational hours or offences issued
- Random Breath Testing (RBT) operations (the number of monthly random breath tests conducted from booze buses and other stationary vehicles)
- Seat belt offences detected (monthly detected seat belt offences)
- Mobile phone offences detected (monthly detected mobile phone

offences).

Mass Media Publicity

- Monthly awareness levels (AdStock) of mass media television advertising with the following themes: speed; fatigue; seat belts; and drink driving.
- Monthly awareness levels (AdStock) of mass media television advertising on all themes.

Change in crash reporting levels from October 2000, associated with changes in the rules for making injury compensation claims following a motor vehicle crash.

Other measures, including legislation and penalty changes. Previous examples include:

- Introduction of the default 50km/h local street speed limit in south-east Queensland in June 1999 (enforcement amnesty period from March to May 1999)
- The Holiday Period Road Safety Trial from December 2001 to end of January 2002
- Introduction of the regional 50km/h local road speed limit from May 2003 (enforcement amnesty period from February to April 2003)
- the increase in speeding penalties from April 2003
- the increase in penalties for use of hand-held mobile phones while driving in December 2003.

It is suggested that the analyses for this Tier remain as recommended in the existing framework. However, in light of the current strategy and action plan, additional examples for the revised framework for this tier could be:

- Roadside Drug Testing Activity
  - Total number of RDTs per month by police region
  - Temporal patterns for testing (time of day, day of week, month of year)
  - Characteristics of motorists tested per month by police region
    - Age
    - Gender
    - Vehicle type
  - Monthly rate of increase in RDTs by police region

- Cycling-related enforcement Activity (including the minimum passing distance road rule)
- Social media inputs and mentions of road safety related themes
- Introduction of double-demerit points for repeat mobile phone use offences
- Changes to Q-Ride.

5.4 Specific Evaluation

A final tier of evaluation recommended for the Queensland Road Safety Strategy is specific evaluation of major program components. The tier 3 models described above measure the general association between measures of specific road safety program activities in a multivariate setting. However, for large complex road safety program elements, specific evaluation of major elements is generally needed for two reasons. First, only specific evaluation can establish the cause and effect relationship between road safety program element implementation and road trauma outcomes with a sufficient degree of scientific rigour. Second, specific evaluation is often needed to establish the measure of road safety program operation that is best related to the outcomes achieved which in turn is fed into the tier 3 models as a key input. For example, specific evaluation of the Queensland speed camera program identified the 6 key measures of program activity listed in the previous section that best predicted crash outcomes.

For these reasons, it is recommended that Queensland Transport continue to commission specific evaluations of key road safety programs implemented as part of the broader road safety strategy. In the past this has included such programs as:

- the mobile speed camera program
- Random Road Watch
- 50km/h default urban speed limits in South East and the rest of Queensland.

It is recommended in the revised framework, as in the existing framework, that Transport and Main Roads continues to commission specific evaluation (where practicable) of programs and legislative changes that support the Strategy goals and objectives.

6 DATA COLLECTION INSTRUMENT/S

In order to assess the traffic safety culture component of the evaluation framework, surveys need to be developed for the key stakeholders and the general community. Specific data collection will be required for the key stakeholders, whereas for the general community the WAVE survey would be adapted to include extra traffic safety culture items. Through the analysis of the literature, the previous WAVE surveys, and other surveys conducted related to road safety attitudes, culture, and behaviour, a number of items and themes have been identified as being useful.

Table 6.1 describes the target population for the surveys, the concepts within traffic safety culture, and the corresponding item/question. The items/questions are further categorised into those that are already included in the TMR WAVE, those that are used by other organisations/jurisdictions, and those that have been developed by the authors. Where necessary items from other surveys have been adapted to apply to the Queensland context (e.g., mobile phone instead of cell phone). It should be noted that these items are suggestions and that further development based on survey development methodologies would need to be applied. Pilot testing and psychometric analysis would also be necessary. The length of any survey should also be considered.

Table 6.1: Target groups, TSC concepts, and suggested survey items

Target group	Concepts	Item
Road safety stakeholders	**What is the state of road safety in Queensland/Australia** How much priority does road safety get/should get? Whose responsibility is it/whose responsibility should it be? How do Queenslanders view road safety now compared to X years ago? What are the key areas that should be targeted in road safety?	**WAVE:** None
		AAA: Please tell us how much of a problem each of the issues below is today compared to 3 years ago (Scale: Much bigger problem today, somewhat bigger problem today, about the same, somewhat smaller problem today, much smaller problem today, don't know) Traffic congestion Aggressive drivers Distracted drivers Drink driving Drivers using drugs
		Developed: What is your role in the safe system? Which aspect of the system is most relevant to your role? Do you see your role as interacting with others in the system? What are the limitations to contributing to the system? (e.g., I am not an engineer so I need to get that expertise elsewhere)
Community: Self	**General attitudes to road safety** Whose responsibility is it? Who does it affect? What are the big issues in road safety? Do you think people get enough	**WAVE: To what extent do you agree with the following statements about road safety? (1 = Agree strongly to 4 = disagree strongly)** Road safety is the responsibility of the entire community My driving affects other road users

	education about road safety?	
		AAA: None
		Developed: To what extent do you agree with the following statements about road safety? (1 = Agree strongly to 4 = disagree strongly) Road safety is my responsibility as a member of the community **How well do you think road safety is performing in Queensland?**

	Attitudes about specific behaviours, risk perceptions Speeding Drink/drug driving Distraction Seatbelts Fatigue Risk perception – general	**WAVE: How risky do you rate the following behaviours?** a) Driving through a stop sign without stopping b) Not wearing a seatbelt c) Travelling UP TO 10 km/hr over the speed limit d) Travelling MORE THAN 10 km/hr over the speed limit e) Driving while tired f) Following another vehicle too closely g) Driving while talking on a hand held mobile phone h) Driving while texting on a hand held mobile phone i) Driving through a give way sign without giving way j) Driving while talking on a hands free mobile phone k) Driving through a red light without stopping l) Driving when you think you may be over the legal alcohol limit m) Driving after having an alcoholic drink (but not enough to be over the legal limit) n) Driving when you are under the influence of illicit drugs o) Overloading the vehicle with too many passengers p) Driving while using a mobile phone or other hand held device to access the internet, take photographs, use maps or use another application

		AAA (adapted): Please tell us how much of a problem each of the issues below is today compared to 3 years ago (Scale: Much bigger problem today, somewhat bigger problem today, about the same, somewhat smaller problem today, much smaller problem today, don't know) Traffic congestion Aggressive drivers Distracted drivers Drink driving Drivers using drugs **AAA (adapted): Where you live, how acceptable would most other people say it is for a driver to...? (**Scale: completely acceptable, somewhat acceptable, somewhat unacceptable, completely unacceptable, don't know) Drive 15 kilometres per hour over the speed limit on a freeway Drive 10 kilometres per hour over the speed limit on a residential street Drive 10 kilometres per hour over the speed limit in an urban area Drive 10 kilometres per hour over the speed limit in a school zone Talk on a hands-free mobile phone while driving Talk on a hand-held mobile phone while driving Type text messages or e-mails while driving Drive when they're so sleepy that they have trouble keeping their eyes open Drive without wearing their seatbelt Drive through a light that just turned red when they could have stopped safely Drive when they think they may have had too much to drink Drive 1 hour after using marijuana Drive after using both marijuana and alcohol

		AAA: How acceptable do you, personally, feel it is for a driver to…? (Scale: completely acceptable, somewhat acceptable, somewhat unacceptable, completely unacceptable, don't know) Drive 15 kilometres per hour over the speed limit on a freeway Drive 10 kilometres per hour over the speed limit on a residential street Drive 10 kilometres per hour over the speed limit in an urban area Drive 10 kilometres per hour over the speed limit in a school zone Talk on a hands-free mobile phone while driving Talk on a hand-held mobile phone while driving Type text messages or e-mails while driving Drive when they're so sleepy that they have trouble keeping their eyes open Drive without wearing their seatbelt Drive through a light that just turned red when they could have stopped safely Drive when they think they may have had too much to drink Drive 1 hour after using marijuana Drive after using both marijuana and alcohol **AAA: Compared to holding a mobile phone in your hand and talking while you were driving, how safe or dangerous do you think it is to talk while driving using a hands-free device?** Scale: Hands-free device is much safer, hands-free device is somewhat safer, they are about the same, hands-free device is somewhat more dangerous, hands-free device is much more dangerous, don't know

		Developed: Where you live, how acceptable would people you are friends with say it is for a driver to...? (Scale: completely acceptable, somewhat acceptable, somewhat unacceptable, completely unacceptable, don't know) Drive 15 miles per hour over the speed limit on a freeway Drive 10 miles per hour over the speed limit on a residential street Drive 10 miles per hour over the speed limit in an urban area Drive 10 miles per hour over the speed limit in a school zone Talk on a hands-free mobile phone while driving Talk on a hand-held mobile phone while driving Type text messages or e-mails while driving Drive when they're so sleepy that they have trouble keeping their eyes open Drive without wearing their seatbelt Drive through a light that just turned red when they could have stopped safely Drive when they think they may have had too much to drink Drive 1 hour after using marijuana Drive after using both marijuana and alcohol
Community: Others	**Perceptions of how road safety is managed and prioritised** Legislation Enforcement Priority areas	**WAVE: How strongly do you support or oppose laws that allow police to impound the vehicle of a driver/rider who has been caught for……..** a) Repeat drink driving b) Repeat disqualified driving c) Repeat unlicensed driving d) Repeat driving an uninsured and unregistered vehicle e) Repeat offences of illegal vehicle modifications f) Repeat drug driving g) "Hooning" offences involving excessive noise or smoke, such as burn outs, donuts, drifting, and other skids / driving stunts

		h) Illegal street racing i) Evading police (failing to stop when directed to do so by a police officer) j) Repeat mobile phone offences **During the last 12 months, have you seen a police vehicle (of any type) patrolling or observing Queensland roads?** **To what extent do you agree or disagree with the following statements…** a) In the past 12 months, the number of police vehicles (of any type) on Queensland roads has increased b) The police are everywhere so I always obey the road rules
		AAA: How strongly do you support or oppose…? (Scale: Support strongly, support somewhat, oppose somewhat, oppose strongly) Having a law against reading, typing, or sending a text message or email while driving Having a law against using a handheld cell phone while driving, for all drivers regardless of their age Having a law against using any type of cell phone while driving, handheld or hands-free, for all drivers regardless of their age Having a law requiring all drivers who have been convicted of DWI to use a device that won't let their car start if they have been drinking, even if it's their first time being convicted of DWI Requiring all new cars to have a built-in technology that won't let the car start if the driver's alcohol level is over the legal limit

		Using cameras to automatically ticket drivers who drive more than 10 mph over the speed limit on freeways Using cameras to automatically ticket drivers who drive more than 10 mph over the speed limit on residential streets Using cameras to automatically ticket drivers who drive more than 10 mph over the speed limit in urban areas Using cameras to automatically ticket drivers who drive more than 10 mph over the speed limit in school zones Using cameras to automatically ticket drivers who run red lights in urban areas Using cameras to automatically ticket drivers who run red lights on residential streets Having a law requiring all motorcycle riders to wear a helmet Having the federal government regulate non-driving-related technologies in cars to make sure they don't distract drivers Having a law making it illegal to drive with more than a certain amount of marijuana in your system Lowering the limit for a driver's blood alcohol concentration from 0.08 to 0.05 g/dL Your state adopting a vision to reduce the number of people killed in accidents to zero

		(Girasek): I would like to see traffic safety included among the president's priorities (Factor: Support increased government attention to traffic safety) (Girasek): The legal blood alcohol limit for driving in the United States is too strict (Factor: Support strict monitoring and control of alcohol-impaired drivers, and other limits on freedom for the sake of traffic safety). (Girasek): If I thought the roads that my family used every day were unsafe, I would contact a public official about them (Factor: Engaged in advancing local traffic safety) (Girasek): Cameras should be installed in more places to enforce traffic safety laws (Factor: Support photo enforcement of traffic safety laws) (Girasek): I would support a $3 increase in the car registration fee to allow my state to do a better job keeping track of where and how crashes occur (Factor: Willing to invest in traffic safety) (Girasek): Before voting on laws that are supposed to make roads safer, politicians should find out what has worked in other places (Factor: Expect evidence-based traffic safety policies) (Girasek): It would be odd for the local PTA to take on road safety as one of its causes (Factor: Believe schools should promote road safety) (Girasek) Item: Police officers who enforce traffic laws are performing a valuable public service (Factor: Support police enforcement and traffic calming measures) (Girasek) I would be interested to know how my state compares to others

		when it comes to the risk of dying in a car crash (Factor: Desire more public access to traffic safety information) (Girasek) I am bothered by car ads that seem to encourage dangerous driving (Factor: Expect the corporate sector to do more about the traffic safety problem)

7 EVALUATION "TEST RUN"

The casualty rate per 100,000 population was modelled for the "before" period January 2001 to December 2003 (the timeframe prior to the previous strategy), using monthly casualty rates. Population figures were based on quarterly estimates from the Australian Bureau of Statistics. Casualty rate trends were then predicted from 2004-2010. Confidence limits (68%) were also estimated for these predicted casualty rates. Casualty rates were calculated based on four different definitions of casualty:

- Police-reported fatalities (Queensland Road Crash Data)
- Police-reported hospitalisations (Queensland Road Crash Data)
- Hospitalised traffic-related casualties[2] (Queensland Hospital Admitted Patients Data)
- Hospitalised traffic-related high threat-to-life casualties[3] (Queensland Hospital Admitted Patient Data)

The first two definitions were consistent with the previous framework with the second two being an example of the recommended additions to the framework. These are included to provide insight into whether the use of the additional data would illuminate any differential trends or impacts of the Strategy.

A time series model (based on state space modelling techniques) was used to model road trauma levels each month before the introduction of the Queensland Road Safety Strategy and forecasts the levels of road trauma that would have been expected to have occurred after the implementation had that strategy not been in place (based on the past trends). Against the forecasts from the resulting model can be plotted the actual road trauma trends in Queensland that occurred after the Queensland Road Safety Strategy was introduced.

7.1 Fatality rate

The first global assessment model estimated was the fatality rate (per 100,000 population) for all of Queensland. The fatal crash risks are shown in Figure 4, in addition, what was observed in the period 2004-2010 (after the implementation of the previous strategy). As shown in the graph, the observed fatality rate trend during the period of post strategy implementation was very similar to that predicted by the model. This indicates that there was not much impact of the strategy on fatality rates following its implementation. Although there was some difference (non-statistically significant lower rate compared to predicted) after mid-2009.

[2] See Section 4.2.2 for selection criteria for traffic-related casualties
[3] High threat to life = ICISS > .941. See Section 4.2.1 for details

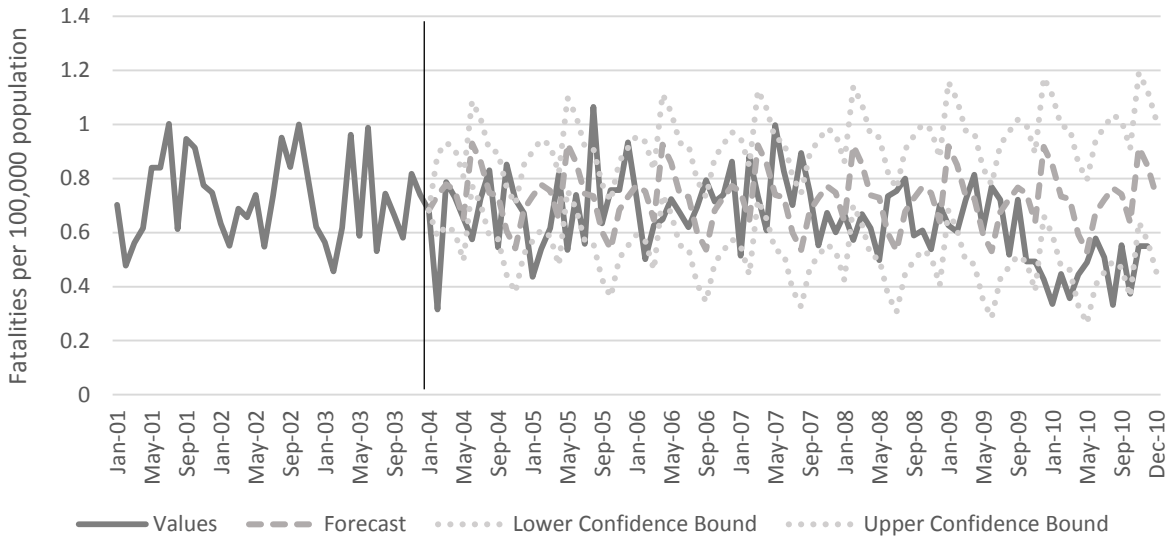

Figure 4: Queensland fatality rate per 100,000 population per month from January 2001 to December 2010

7.2 Police-reported hospitalisation rate

The second global assessment model estimated was the police-reported "hospitalisation" rate (per 100,000 population) for all of Queensland. As shown in, the observed police-reported hospitalisation rate trend during the period of post strategy implementation was very similar to that predicted by the model. This indicates that there was not much immediate impact of the strategy following its implementation. Although there was some difference (non-statistically significant lower rate compared to predicted) after end-2006[4].

[4] There were some issues with the Queensland Road Crash Data coding during the 2006-2008 period that may explain these differences rather than any real change in the hospitalisation rate.

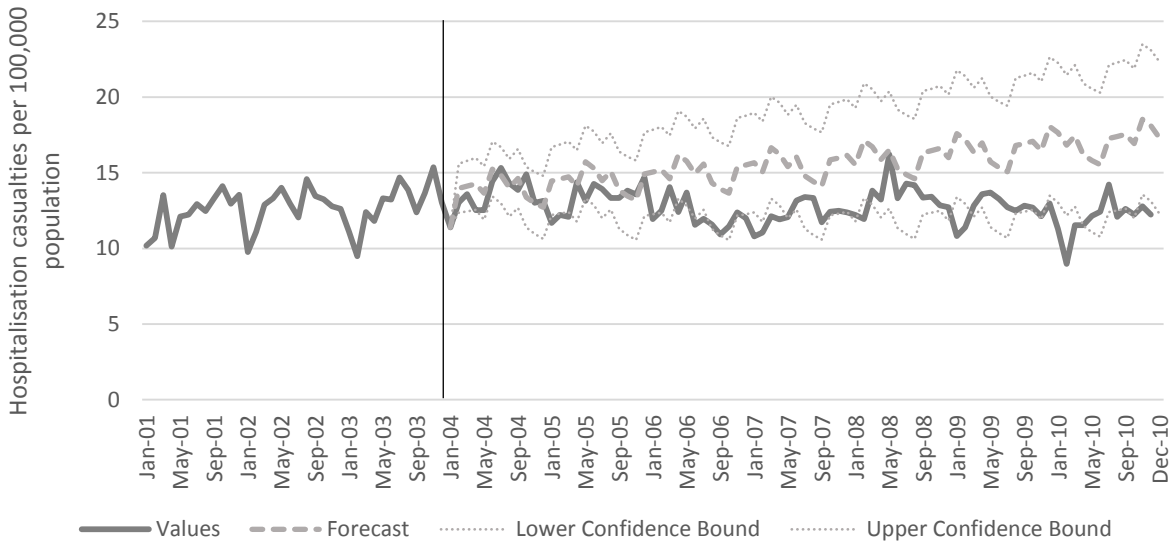

Figure 5: Queensland police-reported hospitalisation rate per 100,000 population per month from January 2001 to December 2010

7.3 Traffic-related hospitalised casualty rate

The third global assessment model estimated was the traffic-related casualties admitted to hospital rate (per 100,000 population) for all of Queensland. As shown in, the observed rate trend during the period of post strategy implementation was very similar to that predicted by the model. This indicates that there was not much impact of the strategy following its implementation. Although there was quite a bit of difference (statistically significant lower rate compared to predicted) after mid-2008.

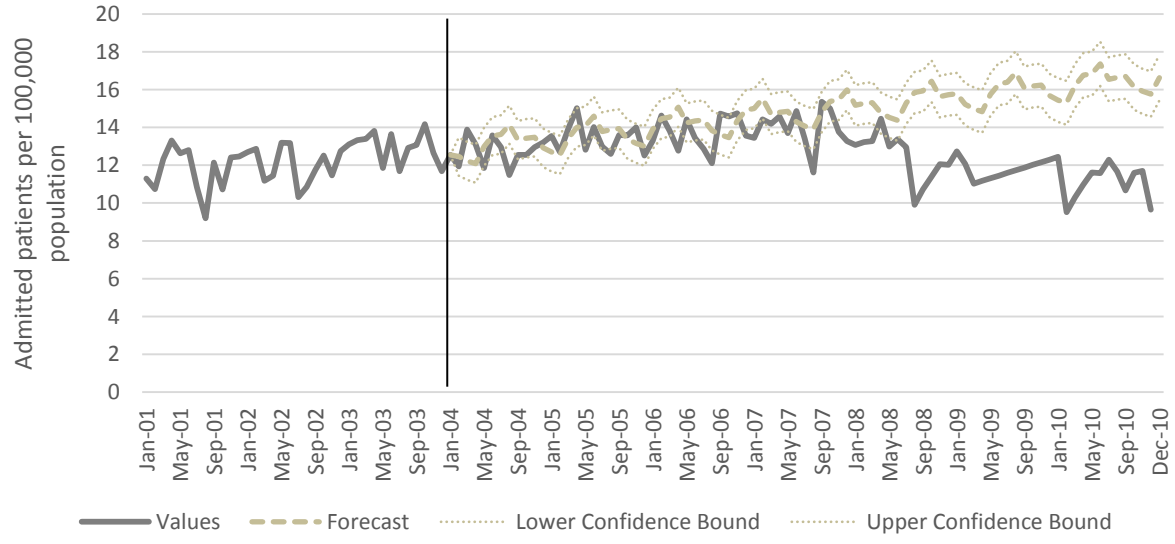

Figure 6: Traffic-related casualties admitted to hospital (both reported and not reported to police) per 100,000 population

7.4 Traffic-related high-threat to life casualty rate

The fourth global assessment model estimated was the traffic-related high threat to life casualty rate (per 100,000 population) for all of Queensland. As shown in, the observed rate trend during the period of post strategy implementation was very similar to that predicted by the model. This indicates that there was not much immediate impact of the strategy following its implementation. Although, there was some difference (non-statistically significant lower rate compared to predicted) after mid-2008.

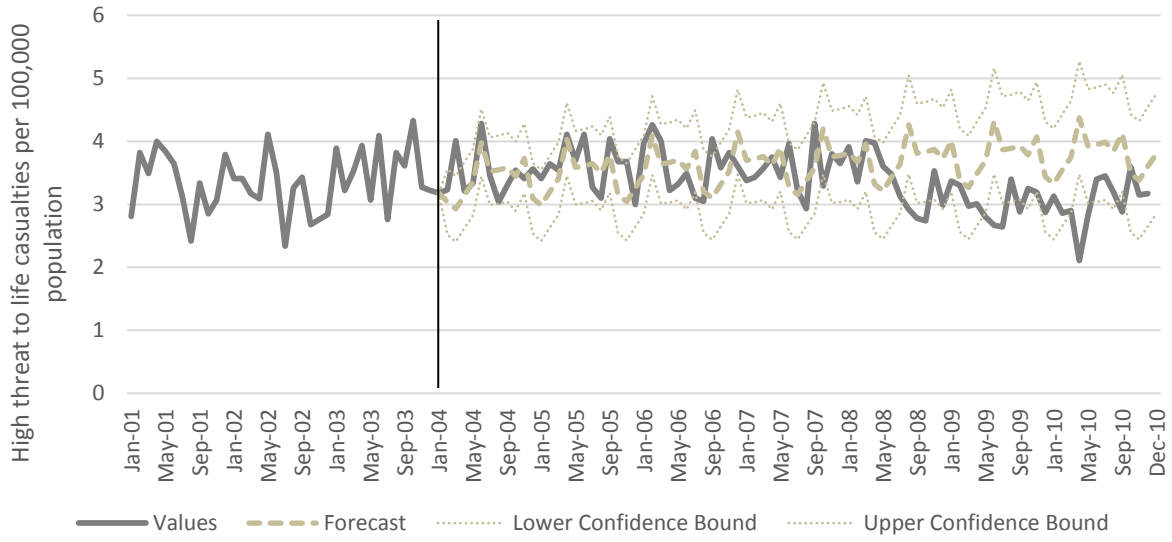

Figure 7: High threat to life traffic-related casualties admitted to hospital (both reported and not reported to police) per 100,000 population

Other elements of MUARC's previous framework test-run were not repeated as it was seen as redundant. Also, for the hospital data the only information for stratification purposes would be road user groups (drivers, passengers, motorcyclists, cyclists, and pedestrians); remoteness location (for the usual residence of the injured person, not the crash location); age; and gender. As described previously, contributing circumstances such as speed, alcohol, or distraction are not available in hospital data. Temporal (time of day and day of week of crash to determine high alcohol hours) and location variables (used to determine speed zone for example) are also not included.

7.5 Police-reported hospitalised versus traffic-related admitted to hospital

As shown in Figure 8, the trend over time for police-reported hospitalised casualty rates was like that of the traffic-related admitted to hospital rates. The only obvious deviation was in the period 2006-2008 which has already been discussed as a data anomaly. While this rate over time is similar, previous research conducted by CARRS-Q has shown that the types of casualties (e.g., cyclists and motorcyclists) vary greatly between the data sources which is disguised by simply observing the overall casualty rates (Watson, Watson, & Vallmuur, 2015).

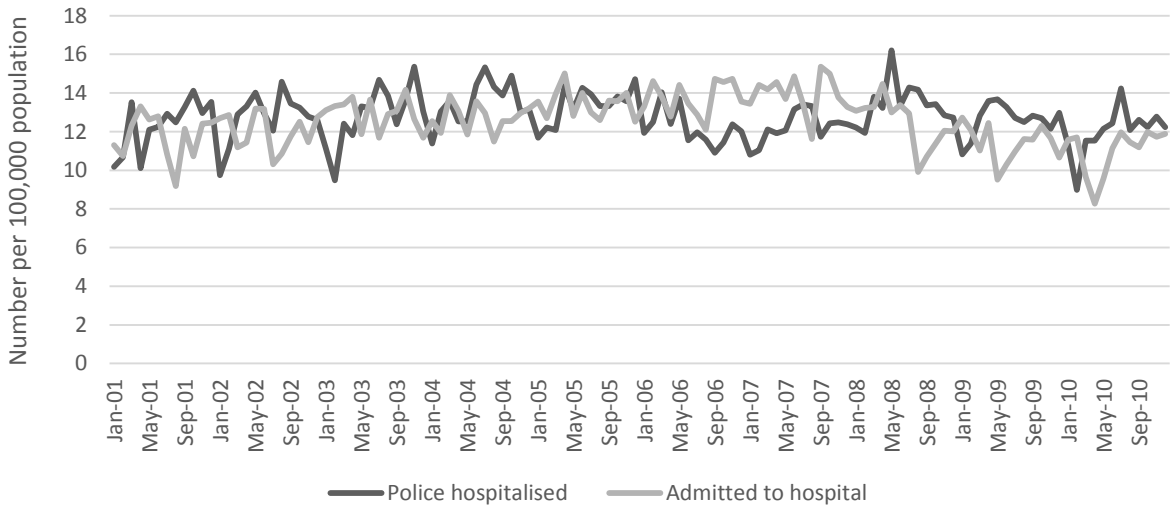

Figure 8: Comparison of trend between police-reported hospitalisations and traffic-related admitted to hospital

8 CONCLUSION AND MILESTONES

Upon review, the existing evaluation framework was found to be a generally comprehensive and reliable model for evaluating contemporary Australian road safety strategies, including the previous Queensland Strategy spanning the years 2004 – 2011. Retaining the multi-tiered GOSPA model as the foundation of the revised evaluation framework is recommended. However, some refinement of the framework was necessary to address some identified gaps and accommodate the new elements and shift in emphases in the Queensland Road Safety Strategy 2015 – 2021.

With Queensland explicating for the first time a need to foster traffic safety culture (TSC), the development of appropriate instruments to monitor the progress toward developing and strengthening a TSC is necessary. The assessment of TSC involves measuring knowledge, attitudes, beliefs, and behaviours, both at stakeholder and community levels, for an overall assessment. Additionally, the existing framework did not address strategy goals, objectives and outcomes explicitly in terms of the Safe System cornerstones, and it is therefore also necessary to develop a suitable scale appropriate to the Safe System approach. Given their close interaction, it is important to account for the TSC - Safe System relationship, which introduces a task of greater complexity than considering them in

isolation. The various dimensions of TSC, its interaction with the Safe System approach, and a range of potential instruments for their measurement have been detailed in this report. The development of appropriate instruments represents a significant step to furthering knowledge and measurement of TSC in Australia.

The current Strategy seeks to move beyond the broad classification of "serious injury" (i.e., generally based on police-reported hospitalisations), as addressed in previous strategies, and able to be monitored through use of the existing framework. The shift in strategy represents an enhanced effort to reduce serious road trauma as well as fatalities, coupled with a more accurate representation of the serious injury burden than has previously been available due to underreporting and a lack of detail in available data. To achieve this, the revised framework seeks to take advantage of improvements in data provision, linkage and analysis in relation to monitoring and addressing serious injuries.

Vulnerable road users including cyclists and motorcyclists are observed to be increasing in number, while their involvement in injury crashes has historically been shown to be significantly underreported through analyses of official reports. The suggested improvements in data provision, linkage and analysis mentioned above will help to provide a more reliable assessment of the magnitude and impact of cyclist and motorcyclist injury. Further, it is recommended that cyclists be included as a specific strata (grouping) for analysis in the revised evaluation framework, having been absent as such from the existing framework.

With the relatively recent expansion of the "fatal four" to the "fatal five" to accommodate distraction as a major contributor to road traffic crashes, it is argued that the revised evaluation framework be tailored to more rigorous assessment and monitoring of its contribution. To assist this closer analysis, it is also recommended to investigate improved means to measure distraction and inattention in their various forms, to more reliably determine their specific role in crashes and assess the effectiveness of relevant countermeasures.

As roadside and random drug testing of drivers has emerged only relatively recently as an enforcement measure, the existing framework contains little specific content related to drug driving. Compared with drink driving, the much lower level of drug driving enforcement and its more targeted nature raises questions about the reliability of information regarding drug driving trends and prevalence. The revised framework seeks to assess drug driving as distinct from drink driving, and it is recommended that more reliable methods for determining and monitoring drug driving rates be investigated.

The existing framework employs a geographical stratification by the eight Queensland police regions. Beyond this, there is a broad and imperfect separation of "urban" from "rural" environments using a speed limit threshold of 80km/h. The adapted framework proposes a refinement of the geographical distinction, based on the Australian Standard Geographic Classification (ASGC) Remoteness Structure to provide separation into five categories (Major Cities; Inner Regional; Outer Regional; Remote; Very Remote). This is considered important because the various Strategy programs and actions are expected to be different in these locations in terms of distribution, coverage and effect.

Finally, there have been some changes in road safety-related regulations and enforcement since development of the existing evaluation framework, including the progression of

drug driving enforcement mentioned above. The minimum passing distance for cyclists and the motorcycle lane filtering laws are two prominent recent examples regarding legislation, but there are also other relatively recent changes with various potential effects that should monitored. The revised evaluation framework will be tailored to allow assessment of these specific changes and related effects.

Table 2 describes the timeframes for implementation of the evaluation framework components as well as the accompanying research/data recommendations to enhance the evaluation framework. Implementation priority is divided into immediate/short-term (within the next 6-12 months), medium term (within 2 years), and longer term (within the life of the Strategy and into the next one). In addition, some recommendations and/or framework components have been categorised as ongoing and others categorised based on the implementation timing of particular programs.

Table 2: Milestone activities and timeframes

Timeframe	Evaluation framework component	Accompanying research and data development
Immediate/short-term (within the next 6-12 months)		Development of additional measures of traffic safety culture (public and stakeholders), including question creation, pilot testing, and preliminary psychometric analysis)
		Improvements to the measures of distraction and its role in crashes be investigated
		Improvements to data to better distinguish between drug and drink driving and to better quantify their relative involvement in crashes be investigated
		Establishment of MOUs for national and state based data linkage
Medium term	The Global Assessment Model (Top Tier Model) can be implemented at the conclusion of the first Action Plan	
	The Second Tier Model (strata) can be implemented at the conclusion of the first Action Plan	

Timeframe	Evaluation framework component	Accompanying research and data development
Longer term	The Global Assessment Model (Top Tier Model) can be implemented at the conclusion of each Action Plan	
	The Second Tier Model (strata) can be implemented at the conclusion of the each Action Plan	
Ongoing	The Global Assessment Model can be used at various time periods after strategy implementation	
	The Second Tier Model can be used at various time periods after strategy implementation	
	Hospital data (not linked) be requested on an ongoing basis to be included in the evaluation model tiers	Opportunities for data linkage and other data improvement strategies be monitored. This includes continued monitoring and support for the Austroads National Data Linkage Project and the Interagency group.
	Rural and remote classification for stratification in the Second and Third Tier be more refined and match the Australian Standard Geographic Classification (ASGC) Remoteness Structure	
	Distraction as a contributor to road crashes should have more emphasis in the Second and Third Tier	
Matched with implementation of	The Third Tier Model can be implemented some	

Timeframe	Evaluation framework component	Accompanying research and data development
programs and action plans (including baseline measures just prior to implementation where possible)	point after program implementation when sufficient post program implementation data has been accumulated	
	Individual program evaluations	

9 REFERENCES

Alsop, J., & Langley, J. (2001). Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accident Analysis & Prevention, 33*(3), 353-359.

Amoros, E., Martin, J.L., Chiron, M., & Laumon, B. (2007). Road crash casualties: characteristics of police injury severity misclassification. *The Journal of Trauma and Acute Care Surgery, 62*(2), 482-490.

Atchley, P., Shi, J. & Yamamoto, T. (2014). Cultural foundations of safety culture: A comparison of traffic safety culture in China, Japan and the United States. *Transportation Research Part F*, *26*, 317-325.

Berry, J.G., Harrison, J.E., & Australian Transport Safety Bureau. (2008). *Serious injury due to land transport accidents, Australia, 2005-06.* Australian Institute of Health and Welfare and the Department of Infrastructure, Transport, Regional Development and Local Government.

Boufous, S., et al., (2008). *Data linkage of hospital and Police crash datasets in NSW.* Sydney: University of New South Wales.

Boyd, J. H., Ferrante, A. M., O'Keefe, C. M., Bass, A. J., Randall, S. M., & Semmens, J. B. (2012). Data linkage infrastructure for cross-jurisdictional health-related research in Australia. *BMC Health Services Research, 12*(1), 480.

Cameron, M.H. (1999). *A method of evaluation based on a strategic planning framework.* Paper presented at the Australasian Road Safety Research, Policing and Education Conference, Canberra.

Chapman, A., & D. Rosman. (2008). *Measuring road crash injury severity in Western Australia using ICISS methodology.* In Insurance Commission of Western Australia Road Safety Forum and Awards, Perth.

Edwards, J., Freeman, J., Soole, D. & Watson, B. (2014). A framework for conceptualising traffic safety culture. *Transportation Research Part F*, *26*, 293–302.

Farmer, C.M. (2003). Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention, 4*(1), 38-44.

Flin, R., Mearns, K., O'Connor, P. & Bryden, R. (2000). Measuring safety climate: identifying the common features. *Safety Science, 34*(1-3), 177-192.

Girasek, D.C. (2012) Towards operationalising and measuring the Traffic Safety Culture construct. *International Journal of Injury Control and Safety Promotion*, *19*(1), 37-46, DOI: 10.1080/17457300.2011.603147.

Haddon, W. (1972). A logical framework for categorizing highway safety phenomena and activity. *The Journal of Trauma*, *12*(3), 193-207.

Hedlund, J. (2007). Improving Traffic Safety Culture in the United States - The Journey Forward: Summary and Synthesis. *AAA Foundation for Traffic Safety,* http://www.aaafoundation.org/pdf/SafetyCultureSummaryAndSynthesis.pdf.

Hofstede, G. (1980). Motivation, leadership, and organizations: Do American theories apply abroad? *Organisational Dynamics, 9*, 42-63.

Hofstede, G., Hofstede, G.J., & Minkov, M. (2010). *Cultures and organizations: Software for the mind (3rd ed.).* New York, NY, USA: McGraw-Hill Professional Publishing.

International Traffic Safety Data and Analysis Group (IRTAD). (2011). *Reporting on Serious Road Traffic Casualties: Combining and using different data sources to improve understanding of non-fatal road traffic crashes.* OECD/ITF.

Langley, J., & Cryer, C. (2012). A consideration of severity is sufficient to focus our prevention efforts. *Injury Prevention, 18*(2), 73-74.

McDonald, G., Davie, G., & Langley, J. (2009). Validity of Police-Reported Information on Injury Severity for Those Hospitalized from Motor Vehicle Traffic Crashes. *Traffic Injury Prevention, 10*(2), 184 - 190.

Moeckli, J., & Lee, J. D. (2007). *The making of driving cultures.* In AAA (Ed.), In *Improving Traffic Safety Culture in the United States: The Journey Forward*. Washington DC: AAA Foundation for Traffic Safety, pp.59-76.

Moore, S. M., Dolansky, M. A., Singh, M., Palmieri, P., & Alemi, F. (2010). *The Systems Thinking Scale.* Unpublished manuscript.

MUARC. (2009). *The development of a framework for road safety strategy evaluation.* Report for Queensland Transport. Melbourne: Monash University Accident Research Centre.

Murray, W., Watson, B., King, M., Pratt S., & Darby, P. (2014). *Applying the Haddon Matrix in the context of work-related safety.* Proceedings of the Occupational Safety in Transport (OSIT) Conference, Gold Coast, 18-19 September 2014. http://ositconference.com/program/2014-proceedings/

Nævestad, T. (2009). Mapping research on high-risk organisations: Arguments for a sociotechnical understanding of safety culture. *Journal of Contingencies and Crisis Management*, 7, 126–136. http://dx.doi.org/10.1111/j.1468-5973.2009.00573.x.

Nævestad, T.-O. & Bjørnskau, T. (2012). How can the safety culture perspective be applied to road traffic?, *Transport Reviews*, *32*(2), 139-154, DOI:10.1080/01441647.2011.628131.

National Centre for Classification in Health. (2004). *The International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification (ICD-10-CM).: Alphabetic index of procedures (ACHI)*. Sydney: National Centre for Classification in Health.

National Collaborative Research Infrastructure Strategy. (2008). *2008 Strategic Roadmap for Australian Research Infrastructure*. Department of Innovation, Industry, Science and Research.

OECD (2016), *Road accidents (indicator).* DOI: 10.1787/2fe1b899-en. Accessed on 22 September 2016. Available: https://data.oecd.org/transport/road-accidents.htm.

Osler, T., et al., (1996). ICISS: an international classification of disease-9 based injury severity score. *The Journal of Trauma and Acute Care Surgery, 41*(3), 380-388.

Queensland Health. (2012). *Queensland Hospital Admitted Patient Data Collection (QHAPDC) Manual* Accessed on 6 October 2016 http://www.health.qld.gov.au/hsu/pdf/manuals/qhapdc12-13/12_13_QHAPDC_FINAL.pdf.

Department of Transport and Main Roads. (2015). *Safer Roads, Safer Queensland: Queensland's Road Safety Strategy 2015-2021.* Queensland Department of Transport and Main Roads.

Rakauskas, M. E., Ward, N. J., & Gerberich, S. G. (2009). Identification of differences between rural and urban safety cultures. *Accident Analysis and Prevention*, *41*, 931–937. http://dx.doi.org/10.1016/j.aap.2009.05.008.

Stephenson, S. C., Langley, J., Henley, G. I., & Harrison, J. E. (2003). *Diagnosis-based injury severity scaling: a method using Australian and New Zealand hospital data coded to ICD-10-AM*. Australian Institute of Health and Welfare.

The Parliament of Victoria Road Safety Committee. (2014). *Inquiry into Serious Injury.*

Ward, N.J., Linkenbach, J., Keller, S.N., & Otto, J. (2010[1]). White Paper on Traffic Safety Culture. White Paper No. 2, *White Papers for: "Toward Zero Deaths: A National Strategy on Highway Safety",* http://safety.transportation.org/doc/web2%20Safety%20Culture%20White%20Paper.pdf.

Ward, H., et al., (2010[2]). *Road Safety Research Report No. 119 Review of Police Road Casualty Injury Severity Classification – A Feasibility Study.* London: Department of Transport.

Watson, A., Watson, B., & Vallmuur, K. (2013). *How accurate is the identification of serious traffic injuries by Police? The concordance between Police and hospital reported traffic injuries.* In Proceedings of the 2013 Australasian Road Safety Research, Policing & Education Conference. Canberra: Australasian College of Road Safety (ACRS).

Watson, A., Watson, B., & Vallmuur, K. (2015). Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accident Analysis & Prevention*, *83*, 18-25.

Wiegmann, D. A., Von Thaden, T. L., & Gibbons, A. M. (2007). A review of safety culture theory and its potential application to traffic safety. In *Improving Traffic Safety Culture in the United States: The Journey Forward*. Washington DC: AAA Foundation for Traffic Safety, pp.77-91.

Williams, A.F., & Haworth, N. (2007). Overcoming barriers to creating a well-functioning safety culture: A comparison of Australia and the United States. In *Improving Traffic Safety Culture in the United States: The Journey Forward*. Washington DC: AAA Foundation for Traffic Safety, pp. 77-91.

World Health Organization. (2010). *Data systems: a road safety manual for decision-makers and practitioners.* Geneva: World Health Organization.