MONASH
UNIVERSITY

RESPONSE BY MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE

TO

"Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns" (White, Walker, Glonek and Burns, November 2000)

Max Cameron
Stuart Newstead

MONASH UNIVERSITY
ACCIDENT RESEARCH CENTRE

NOVEMBER 2000

MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE
REPORT DOCUMENTATION PAGE

Report No.	Date	ISBN	Pages
177	November 2000	0 7326 1476 7	xi + 30

Title and sub-title:

Response by Monash University Accident Research Centre to *"Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns"* (White, Walker, Glonek and Burns, November 2000)

Author(s)
M.H. Cameron and S.V. Newstead

Type of Report & Period Covered:
General, 1992-2000

Sponsoring Organisation(s):
This project was funded through the Centre's Baseline Research Program for which grants have been received from:
Department of Justice
Roads Corporation (VicRoads)
Royal Automobile Club of Victoria (RACV) Ltd
Transport Accident Commission (TAC)

Abstract:

This document is the response by Monash University Accident Research Centre (MUARC) to an investigation of some specific MUARC research. The investigation has been published by the South Australian Department of Transport (White, Walker, Glonek and Burns 2000). It relates principally to MUARC's research on traffic enforcement and road safety advertising in Victoria. From re-analysis of data used in MUARC report no. 74, White et al have concluded that the estimates of crash and financial savings attributable to the TAC-funded countermeasures are not supportable. They have also concluded that the re-analysis has failed to support the claims of MUARC report no. 52 concerning the crash reductions that can be achieved through high levels of TAC-funded road safety television advertising.

In this document, MUARC responds to White et al and reaches the following conclusions:

1. Scientific evaluations conducted by MUARC have shown substantial reductions in road trauma in Victoria due to increased random breath testing using "booze buses" and the new speed camera program, each supported by TAC advertising.
2. The statistical models of monthly casualty crashes as functions of enforcement, advertising and socio-economic factors, developed in MUARC report no. 52, are sound. They have been tested by MUARC and by White et al and have been found to be satisfactory. White et al's investigations have provided additional evidence of the relationship between the TAC speed-related advertising and crashes.
3. The estimates of the points of diminishing returns of levels of TAC drink-driving and speed-related advertising, originally provided in MUARC report no. 52 based on the statistical models in that report, are sound. The economic analysis of advertising levels, which was based on the coefficients of the advertising variables in the statistical models, has not been questioned.
4. White et al's re-analysis of the data used in MUARC report no. 74 is not relevant to report no. 52 because of important differences in their objectives, the types of crashes analysed, the time periods covered, the treatment of levels of speed-related advertising, the inclusion of car-based random breath testing, and the assumptions made and subsequently tested.
5. White et al's so-called parsimonious three-factor model of crash variations in Victoria was based on data dredging and cannot be considered to be a valid alternative to MUARC's models.
6. White et al's test of the quantitative relationships between crashes and the enforcement and advertising variables is not valid. It is not an adequate test of the presence or absence of quantitative relationships.

Key Words:
Response; Reanalysis; Evaluation; Trends; Road Crashes; Casualties; Transport Accident Commission; Television Advertising; Mass Media; Campaigns; Enforcement; Speeding; Drink Driving; Unemployment; Leading Index of Economic Indicators; Business Cycle; Econometric Analysis; Victoria; Australia; Monash

Monash University Accident Research Centre,
Wellington Road, Clayton, Victoria, 3800, Australia.
Telephone: +61 3 9905 4371, Fax: +61 3 9905 4363

PREFACE

This document was prepared at the request of the Committee responsible for the Baseline Research Program of Monash University Accident Research Centre (MUARC). During 2000, the members of the Committee and the organisations they represented were:

- Mr Peter Eynaud, Department of Justice
- Dr Ken Ogden, Royal Automobile Club of Victoria (RACV) Ltd.
- Mr David Healy, Transport Accident Commission
- Superintendent Bob Wylie, Victoria Police
- Dr Jeff Potter, VicRoads
- Professor Claes Tingvall, Director, MUARC
- Professor Tom Triggs, Deputy Director, MUARC

This document has the full support of the members of the Baseline Research Program Committee.

MUARC believes that this document provides a full response to the *"Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns"* by White, Walker, Glonek and Burns (November 2000). No further comments on this issue will be made by MUARC.

Table of Contents

RESPONSE BY MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE

TO

"Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns"
(White, Walker, Glonek and Burns, November 2000)

EXECUTIVE SUMMARY

INTRODUCTION

This document is a summary of the response by Monash University Accident Research Centre (MUARC) to an investigation of some specific MUARC research. MUARC welcomes reviews of its work because of the critical importance of many of MUARC's results and conclusions. MUARC endeavours to conduct its research to the highest possible scientific standards, while recognising the need to provide results and advice in a timely manner. When necessary, MUARC always qualifies any of its work which is less than definitive.

The investigation relates principally to MUARC's research on traffic enforcement and road safety advertising in Victoria. MUARC is independent of any policy considerations which favour or disfavour various strategic approaches to the use of these measures to achieve road safety goals. MUARC's aim, within the constraint of the resources available to it, is to provide objective information on the effectiveness and cost-effectiveness of these measures in the broader context of factors affecting road trauma. Scientific principles have always been applied by MUARC to achieve this aim.

THE RE-INVESTIGATION PROCESS

In April 1997, the South Australian Office of Road Safety wrote to MUARC requesting that the data used for MUARC report no. 52 (RN52), *Evaluation of Transport Accident Commission road safety television advertising* (Cameron, Haworth, Oxley, Newstead and Le 1993), be supplied to the Office so that testing of MUARC's statistical models could be carried out.

During 1998-2000, MUARC was provided with eight draft reports, totalling 889 pages, produced during the investigation by the Office of Road Safety. The first two reports focused on the data used in MUARC report RN52. The third and subsequent reports focused on the data from MUARC report RN74, *Modelling of some major factors influencing road trauma trends in Victoria 1989-93* (Newstead, Cameron, Gantzer and Vulcan 1995).

A paper by White, Walker, Glonek and Burns, *Re-investigation of the Effectiveness of the Victorian TAC's Road Safety Campaigns,* has been published (White et al 2000a). The final report on the investigation has also been released (White et al 2000b). This document is a summary of MUARC's response to the paper and report. MUARC's response also refers to material in the earlier Office of Road Safety reports. Interested readers are strongly advised to consult MUARC's full response for complete understanding of the many issues involved.

MUARC RESEARCH ON ENFORCEMENT AND ADVERTISING

White et al (2000b) claim that "Although a number of different statistical techniques were employed by MUARC in their earlier evaluations [of enforcement and media campaigns], multiple linear regression became the favoured technique in their later work".

MUARC has completed thirteen substantial studies involving the analysis of real crash data in relation to Police enforcement, road safety advertising and sometimes additional factors. Details are given in the full response. Five of the MUARC studies (or parts thereof) have used multiple regression time series analysis to link crashes with road safety program measures and other factors. The remaining ten studies and part-studies have all been quasi-experimental evaluations of the impacts of the programs. The quasi-experimental time series evaluations were not critically dependent on the need to develop crash models including all influential factors in order to reach their conclusions. MUARC has used a mixture of the study methods over the last decade.

Multiple regression [time series analysis] has not been the favoured statistical technique used by MUARC in this area. It has not been used in any study which MUARC has regarded as being a scientific evaluation of enforcement and/or advertising and has not needed to qualify the results. Multiple regression time series analysis has been used only in studies either attempting to represent the underlying mechanisms of road safety programs or studies attempting to consolidate previous findings.

White et al (2000a) claim that "Much of the MUARC research was brought together in a report by Newstead, Cameron, Gantzer and Vulcan (1995)", ie. RN74, and that this is "representative of the MUARC research that has been influential in shaping government policy on road safety television advertising levels" (White et al 2000b).

RN74 is not representative of the MUARC evaluation research on enforcement and advertising. RN74 is not even representative of the research in RN52, which may have been influential in government policy. Details of important differences between RN52 and RN74 are given later.

MUARC has also studied the trends in the Victorian annual road toll at the macroscopic level. These studies have indicated that the decrease in the Victorian road toll during 1990 and subsequent years was not consistent with expected trends and was a greater reduction than expected from pre-existing trends. These findings do not constitute an evaluation of the 1990 initiatives, but the findings do provide evidence of a real change in road safety in Victoria during the 1990's.

MUARC CONSULTANCY ADVICE

MUARC has been commissioned to provide road safety policy advice to jurisdictions outside Victoria. On occasions, this advice has been based on MUARC's evaluations of Victorian programs and/or the expert opinion of the nominated consultants.

In 1996, MUARC was commissioned by the South Australian Office of Road Safety to prepare the report, *Possibility of adapting some road safety measures successfully applied in Victoria to South Australia* (Vulcan, Cameron, Mullan and Dyte 1996), RN102. The report recommended, *inter alia*, "that resources be allocated to double the exposure of television advertisements which support the speed camera and random breath testing programs".

Referring to this recommendation, White et al (2000a,b) state that "The main aim of the project reported here [the re-investigation] is to check the soundness of the advice concerning the effectiveness of high levels of television advertising, through re-analyses of the data originally analysed by Newstead et al (1995)".

The consultancy report, RN102, makes no reference to Newstead et al (1995), RN74. Reference is made to RN52, but this was not the only basis on which the consultants advised their recommendation. RN102 included a summary of the RN38 and RN42 evaluations which indicated the substantial reductions in severe crashes associated with the Victorian enforcement and advertising programs aimed at drink-driving and speeding. RN102 also noted that the level of road safety television advertising in Adelaide during 1993-1995 was less than half the level in Melbourne during the same years.

In all of MUARC's consultancy advice to other jurisdictions related to road safety advertising in support of enforcement, reference has been made to the evaluation studies RN38 and RN42, as well as to RN52. This has been considered appropriate, given that RN52 includes a number of assumptions and qualifications on which it was based.

IMPORTANT DIFFERENCES BETWEEN RN52 AND RN74

White et al (2000b) state "It is considered that two selected RN74 analyses are representative of the MUARC research that has been influential in shaping government policy on road safety television advertising levels". It is possible that RN52 may have been influential, if its stated assumptions were accepted by its readers. However, it is <u>not</u> correct that the research in RN74 is representative of that in RN52.

There are important differences between RN52 and RN74 in their objectives, the type of crashes analysed, the time periods covered, the treatment of levels of speed-related advertising, and the inclusion of car-based random breath testing. The economic analysis in RN52 was based on models of all casualty crashes (not just serious casualty crashes), the period 1983-1991 for the LAH[1] casualty crashes (not 1983-1992; this period was used for the HAH crashes), the levels of speed-related advertising were measured in TARPs (not Adstock; this was used only for drink-driving advertising), and the random breath test data included car-based tests (not just the bus-based tests included in the model in RN74).

The aim of RN52 was to provide advice on the effectiveness of TAC advertising; in particular the relationship between the level of advertising investment and reductions in injury costs. It was expected that this urgent study would require assumptions to be made.

[1] LAH, the "low alcohol hours" of the week (ie. Monday-Thursday 6am to 6pm, Friday 6am to 4pm, Saturday 8am to 2pm, Sunday 10am to 4pm), are those periods when the percentage of drivers killed or admitted to hospital with a blood alcohol content exceeding 0.05%, was below 4%. HAH, the "high alcohol hours", are the converse of these periods, during which about 38% of driver serious casualties had blood alcohol content exceeding 0.05% (Harrison 1990).

The modelling of monthly casualty crash variations was based on assumptions about the functional form of the relationships, specific explanatory variables, and the correlations between these variables. RN52 clearly stated the assumptions. The reader was cautioned that assumptions had been made in a total of six places in the report. During the project there was limited diagnostic testing of the models in the time available. However, the tests done and subsequent testing by White et al have found that the models were satisfactory.

The aim of RN74 was not to evaluate the enforcement and advertising programs, it was to combine the results of evaluations of Victorian road safety programs and information on the influence of two socio-economic factors (unemployment rates and alcohol sales) to assess the contribution of each program/factor to Victorian road trauma trends during the early 1990's. RN74 did not have a specific focus on the advertising component of the Victoria's road safety initiatives, unlike RN52.

The narrower focus of the crashes considered for the analysis in RN74, compared with RN52, is the most critical difference between these two studies. White et al (2000b) state that "It is considered unlikely that the main points to be made in this report would be different if any if any other MUARC analyses had been selected for re-investigation". MUARC's response is that it is not unlikely, if RN52 had been the focus, contrary to White et al's (2000b) opinion.

DATA DREDGING

White et al (2000a) claim that "The methods employed for the selection of variables in the MUARC modelling process can be described as 'data dredging' ". To support this claim, White et al (2000b) quote principally from RN29, *Linking economic activity, road safety countermeasures and other factors with the Victorian road toll* (July 1992).

RN29 was MUARC's initial analysis of factors linked to road trauma trends in Victoria. It examined relationships with monthly fatalities only and considered the potential explanatory variables within a conceptual framework used to minimise any spurious selection due to chance. RN29 was not considered to be a scientific evaluation of the factors. It did suggest important influential factors, such as those representing economic conditions, which needed to be taken into account in MUARC's subsequent evaluations of Victoria's road safety initiatives.

The quasi-experimental time series evaluations of the RBT "booze bus" initiative (RN38) and the increased speed cameras (RN42), with the supporting publicity in each case, needed to take into account the different trends in vehicle travel and/or unemployment rates in Victoria and NSW to ensure the integrity of the evaluation design. Unemployment rate was chosen for non-arbitrary reasons associated with the specificity of the data and theories about a causal role. Separate rates were available for both the metropolitan and rural areas of Victoria and NSW. Unemployment rate was also considered to represent variations in discretionary, higher-risk travel. Details are given in MUARC's full response.

The multiple regression time series analyses (RN52 and RN74) continued to use unemployment rates as an explanatory factor without consideration of alternative economic variables. Initially the models used TARPs to represent road safety television advertising as an explanatory factor, but when MUARC became aware of Adstock and its conceptually

better basis for a link with awareness and hence potentially with crashes, Adstock became the preferred measure.

This process of evolution of the choice of covariates and explanatory factors for inclusion in MUARC's quasi-experimental evaluations and crash modelling analyses, respectively, did not constitute "data dredging" as implied by White et al (2000a,b). The factors were chosen with careful attention to avoiding spurious inclusion and on the basis of reasoned consideration of their possible causality and potential explanatory role.

White et al (2000b) have proposed an alternative to unemployment rate as a measure of economic activity (outlined below). They have apparently considered a range of indicators for this role, with a range of lags, during their investigation. In their 1999 re-investigation reports they proposed that monthly unemployment rate, brought forward by 12 months, would be adequate in providing the sole explanation for the trends in serious casualty crashes during 1983-92.

MUARC does not consider it is appropriate to use "data dredging" to select a factor to explain the variation in a road crash data series. If sufficient factors, with a variety of leads and lags, are considered the analyst is almost certain to find a factor which explains the series well. However this approach runs the danger that the selected factor has no causal basis for the explanation, and that the apparent relationship is spurious. The use of a conceptual or theoretical model to select the factors for consideration is important to minimise this danger.

WHITE ET AL'S ALTERNATIVE INDEX OF ECONOMIC ACTIVITY

White et al (2000a,b) have proposed the Leading Index (LI) of Economic Indicators for inclusion in the models in RN74 instead of unemployment rate. MUARC questions the process through which White et al have chosen and used this variable as well as the conclusions they reach.

White et al (2000b) present arguments for the choice of LI which include "The peak in the Leading Index occurs at much the same time as the peak in all (ie, LAH + HAH) casualty crashes". MUARC cautions that the inferences which are made from pair-wise comparisons of time series variables can be very misleading when the relationships between crashes and other factors are truly multivariate. The peak in the crash series may change after adjustment for the influence of other factors.

White et al (2000b) state "[A] reason for selecting this measure was, admittedly, because it peaked at much the same time as the peak in crash numbers. In that respect the measure is a product of 'data dredging', … "

White et al (2000b) also state "Because the month-to-month variation in the Leading Index was shown … to be unrelated to the month-to-month variation in crash numbers, a smoother version of the Leading Index was created by taking the 12-month centred-moving-average (CMA 12)." These are remarkable statements about the choice and use of an economic indicator to replace the one which MUARC has found satisfactory. MUARC questions how future LI values (up to six months ahead) could be causally related to the crashes in a specific month?

When White et al (2000b) included the CMA 12 Leading Index in the RN74 models, the magnitudes of the estimated coefficients of the enforcement and advertising variables, and their statistical significance levels, were reduced. MUARC considers this to be an inevitable outcome of a process whereby an alternative economic indicator has been chosen on the basis of its coinciding peak and general correlation with the crash series.

Despite MUARC's concerns, in White et al's (2000b) LAH serious casualty crash model which included the CMA 12 Leading Index, both the speed-related enforcement and advertising variables were statistically significant. These findings are evidence of a link with crashes even after the explanation associated with White et al's CMA 12 Leading Index is taken into account.

White et al (2000b) also considered a model of all (LAH plus HAH) serious casualty crashes, with the CMA 12 Leading Index and the speed-related enforcement and advertising variables as explanatory factors. The speed-related advertising was found to be statistically significant, however the enforcement variable was not.

WHITE ET AL'S THREE-FACTOR MODEL

White et al (2000a,b) have proposed that a three-factor model incorporating (1) linear trend, (2) seasonality, and (3) the CMA 12 Leading Index would be adequate to explain the monthly variations in all Melbourne serious casualty crashes during 1983-1992. They discount the contribution of the speed-related advertising to explaining variations in the same crash series, saying that "speed advertising made only a weak contribution". In fact, White et al (2000b) had shown that the contribution of the speed-related advertising was highly statistically significant ($p = 0.006$).

White et al (2000b) have presented the results of fitting an additive model of their three factors to the monthly crashes, rather than a multiplicative model. The multiplicative model form has been used throughout MUARC's multiple regression time series analysis. The additive model cannot be claimed to be more parsimonious than a multiplicative model. White et al (2000b) appear to suggest that an additive model is simple. The multiplicative functional form used by MUARC ensures that the number of crashes predicted by the model cannot be negative. An additive model would not necessarily meet this constraint and could lead to incorrect conclusions about the significance of factors included in the model.

MUARC questions the change in functional form of the crash model for this analysis. If the multiplicative form had been retained by White et al (2000b), their own analysis shows that, had the speed-related advertising been considered for inclusion in the three-factor model, it would have been statistically significant. The three-factor model cannot be described as parsimonious, because it does not take into account at least one factor (ie. speed-related advertising) known also to be associated with crashes.

TIMING OF TURN-AROUND(S) IN VICTORIAN CRASHES

White et al (2000a,b) state that "Reports from some Victorian and overseas road safety agencies give the clear impression that the agencies believed that the TAC-funded

enforcement and advertising campaigns were largely, if not entirely, responsible for *halting the increase* in crash numbers and for *initiating their decline*."

White et al (2000b) state that these "misunderstandings … are in no way attributable to the work of the MUARC researchers". The most that MUARC has claimed is that a number of road safety measures and other factors have *contributed* to the reductions in road trauma in Victoria during 1990 and later years.

White et al (2000a,b) claim that the start of the decline in serious casualty crashes preceded the launch of the speeding and drink-driving enforcement and advertising campaigns, and that this casts doubt on their causal role. This analysis may be valid if a single factor (ie. the relevant enforcement or advertising) had been responsible for the trends in the crashes during 1983-92. However, MUARC has identified many factors which are associated with the trends in crashes during this period. The single-factor comparisons made by White et al (2000a,b) ignore the simultaneous effects the other factors have had on the observed crash series.

RN52 shows that the peak in casualty crashes appears to be in 1989 for both HAH and LAH. However the peak in serious casualty crashes appears to be in 1988. MUARC has offered an explanation for the different trends in serious casualty crashes compared with all casualty crashes (of which serious casualty crashes were about one-third); details are given in MUARC's full response. It is emphasised that MUARC's research has been confined to assessing factors which contributed to the reductions in road trauma in Victoria during the 1990's. MUARC has not evaluated the factors which may have been responsible for the turnaround in crashes of each level of severity prior to 1990.

TESTS OF QUANTITATIVE RELATIONSHIPS

White et al (2000a,b) suggest that RN52 and RN74 claimed that quantitative relationships between TAC advertising levels and crashes had been identified. They present the results of tests based on the data in RN74 which they claim fail to provide evidence of any quantitative relationships.

The tests involved defining a dichotomous (0, 1) variable for each of the four speeding and drink-driving enforcement and advertising measures and then including each pair (dichotomous and raw variable) together in the original multiple regression analysis of the relevant serious casualty crashes. In each case the formerly statistically significant enforcement and advertising variables (in raw form) became non-significant and the dichotomous variables were also non-significant in every case. White et al (2000a,b) argued that, if the relationships had been truly quantitative, the raw measures should have retained their statistical significance. Failure to do so constituted a lack of evidence of a quantitative relationship, in White et al's opinion.

MUARC does <u>not</u> agree that the test performed by White et al (2000a,b) is an adequate test for the presence or absence of any quantitative relationships. It is known that if two highly correlated variables are included together in a multiple regression, then statistically meaningless results will occur. In their 1999 re-investigation reports, White et al outlined the same tests on the RN74 data. The correlations between the pairs of raw and dichotomised variables were greater than 0.99 in three out of the four cases. They stated "there is little point

in even attempting such an analysis if the correlations between the dichotomous and quantitative versions of the same variable are so high that statistical problems of collinearity would necessarily be introduced into the regression analysis. This could be the case for correlations above 0.90 (Tabachnik and Fidell, p. 96) and would almost certainly be the case for correlations of 0.99 or above".

MUARC is surprised that, although they had clearly recognised the problem of high collinearity in the context of their test, White et al (2000b) decided to proceed with it. MUARC rejects the notion that the test is a valid test of the quantitative relationships. MUARC considers that the presence or absence of quantitative relationships linking crashes with speed-related advertising and with drink-driving enforcement and advertising has not been adequately tested by White et al (2000b).

However, the 1999 re-investigation reports include additional analyses which indicate the presence of quantitative effects of the speed-related advertising. In the reports, White et al drew a distinction between the macro- and micro-level effects of the advertising and the enforcement on crashes. The micro-level effects could be described as quantitative.

White et al analysed LAH serious casualty crashes during the period from August 1990 to December 1992, after the speed camera program commenced full operation. They found negative correlations with the monthly speed camera tickets ($p = 0.097$) and speed-related Adstock ($p = 0.038$), the latter being statistically significant. White et al (2000b) dismiss this earlier analysis as not being appropriate and do not report it. Nevertheless, the analysis did find statistically significant evidence of micro-level effects of the speed-related advertising.

In summary, White et al's (2000a,b) test of the presence or absence of quantitative relationships was not considered valid because of high collinearity problems associated with three of the four enforcement and advertising measures under consideration. An alternative analysis by White et al indicated that there is a quantitative relationship between speed-related advertising and crashes.

STARTING MONTH OF DRINK-DRIVING ADVERTISING

White et al (2000b) have criticised MUARC researchers for using November 1989 as the start date of drink-driving advertising for the analysis behind RN52 and RN74, arguing that all other documents have used mid-December 1989 as the start of TAC road safety advertising.

MUARC had non-arbitrary reasons for the inclusion of levels of drink-driving advertising during November 1989 in the data used to develop the statistical models in RN52 and RN74. Details are given in the full response. MUARC's inability to include advertising levels prior to November 1989 was due to the absence of relevant information in useable form.

MUARC disagrees with White et al (2000b) about the criticality of the November 1989 data to MUARC's modelling results. MUARC did not find this to be the case when it re-analysed the model developed in RN52, which is the only MUARC report relevant to the question of road safety advertising levels.

WHITE ET AL'S CONCLUSIONS AND RECOMMENDATIONS

White et al (2000b) conclude "From the re-analyses of the data of Newstead et al (1995) [RN74] it is concluded that the … estimates of crash and financial savings attributable to the TAC-funded countermeasures are not supportable".

MUARC responds that RN74 is not representative of the MUARC research on enforcement and advertising in Victoria during the 1990's. Scientific evaluation studies, especially RN38 and RN42, have shown substantial reductions in road trauma due to the RBT "booze buses" and the new speed cameras, each supported by TAC advertising.

White et al (2000b) also conclude "More particularly, the re-analyses [of RN74 data] have failed to support the claims of Cameron et al (1993) [RN52] concerning the crash reductions that can be achieved through high levels of TAC-funded road safety TV advertising".

MUARC responds that the data and analysis in RN74 is not representative of that in RN52. There were important differences in the objectives, data analysed, and assumptions of these two studies. RN52 provided estimates of the point of diminishing returns for TAC advertising, subject to stated assumptions of the analysis, whereas RN74 did not. The assumptions of RN52 have been found to be satisfactory. White et al's findings from the re-analysis of the data in RN74 are not relevant to RN52.

White et al (2000b) recommend that "The developers of an innovative road crash countermeasure should always first consider the possibility of implementing the countermeasure in such a way that it can be evaluated experimentally". MUARC supports this recommendation very strongly. However this has seldom happened in Victoria and MUARC has been forced to use quasi-experimental designs in its evaluations of the enforcement and advertising programs. MUARC recognises that evaluation on this basis is not ideal, but represents the best available approach in non-experimental settings.

White et al (2000b) also recommend that "MUARC consultancy advice in favour of very high levels of road safety advertising should not be taken into consideration when determining appropriate levels of such advertising". MUARC re-iterates a number of points with respect to its consultancy advice and the link between TAC road safety advertising and crashes:

- RN38 and RN42 provided scientific evidence of reductions in severe crashes due to:
  - the RBT "booze buses", supported by TAC drink-driving advertising
  - the new speed cameras, supported by TAC speed-related advertising

- MUARC's macro-level trend analysis showed that the decrease in the Victorian road toll during 1990 and subsequent years was a greater reduction than that expected from pre-existing trends

- RN52 estimated the points of diminishing returns, under stated assumptions, for levels of drink-driving and speed-related TAC television advertising, respectively

- the assumptions made by MUARC in RN52 have been tested and found to be satisfactory

- the advice provided to the South Australian Office of Road Safety in RN102 made reference to RN38, RN42 and RN52 (but not RN74) and to data which showed that the level of road safety advertising in Adelaide was less than half the level in Melbourne

- the advice was not dependent on the findings in RN52 (had the advice been based on the points of diminishing returns, the recommendation would have been to more than double the level of advertising in South Australia)

- the statistical models of monthly crash variations developed in RN52 were not based on "data dredging" to find enforcement, advertising and socio-economic variables, which then may have had only spurious relationships with the crashes, to include in the models

- when White et al's (2000b) CMA 12 Leading Index replaced the unemployment rate in the RN74's LAH serious casualty crash model, the speed-related enforcement and advertising variables remained as statistically significant factors explaining the crash variations

- speed-related advertising was a statistically significant factor in White et al's (2000b) CMA 12 Leading Index model of monthly serious casualty crashes during all times of the week

- White et al's (2000b) test of whether the enforcement and advertising variables in the RN74 models were quantitatively related to the crash variations is not valid and does not establish the absence of quantitative relationships

- White et al's (1999) earlier investigation of the micro-level effects on LAH serious casualty crashes found that the speed-related advertising was statistically significant

These points indicate that the basis for MUARC's consultancy advice is sound. Based on White et al's investigations, the evidence in favour of the effectiveness of TAC speed-related advertising supporting the Victorian speed camera program has been strengthened. The evidence for the effects of the drink-driving advertising supporting Victoria's RBT program relies on RN38 and on relevant parts of RN52 (which White et al have tested and found satisfactory).

MUARC'S CONCLUSIONS

1. Scientific evaluations conducted by MUARC have shown substantial reductions in road trauma in Victoria due to increased random breath testing using "booze buses" and the new speed camera program, each supported by TAC advertising.

2. The statistical models of monthly casualty crashes as functions of enforcement, advertising and socio-economic factors, developed in RN52, are sound. They have been tested by MUARC and by White et al and have been found to be satisfactory. White et al's investigations have provided additional evidence of the relationship between the TAC speed-related advertising and crashes.

3. The estimates of the points of diminishing returns of levels of TAC drink-driving and speed-related advertising, originally provided in RN52 based on the statistical models in

that report, are sound. The economic analysis of advertising levels, which was based on the coefficients of the advertising variables in the statistical models, has not been questioned.

4. White et al's re-analysis of the data used in RN74 is not relevant to RN52 because of important differences in their objectives, the types of crashes analysed, the time periods covered, the treatment of levels of speed-related advertising, the inclusion of car-based random breath testing, and the assumptions made and subsequently tested.

5. White et al's so-called parsimonious three-factor model of crash variations in Victoria was based on data dredging and cannot be considered to be a valid alternative to MUARC's models.

6. White et al's test of the quantitative relationships between crashes and the enforcement and advertising variables is not valid. It is not an adequate test of the presence or absence of quantitative relationships.

REFERENCES

CAMERON, MH, HAWORTH, N, OXLEY, J, NEWSTEAD, SV, and LE, T (1993), "Evaluation of Transport Accident Commission road safety television advertising". Report No. 52, Monash University Accident Research Centre.

HARRISON, WA (1990), "Update of alcohol times as a surrogate measure of alcohol-involvement in accidents". Research Note, Monash University Accident Research Centre.

NEWSTEAD, SV, CAMERON, MH, GANTZER, S, and VULCAN, P (1995), "Modelling of some major factors influencing road trauma trends in Victoria 1989-93". Report No. 74, Monash University Accident Research Centre.

VULCAN, AP, CAMERON, MH, MULLAN, N, and DYTE, D (1996), "Possibility of adapting some road safety measures successfully applied in Victoria to South Australia". Report No. 102, Monash University Accident Research Centre.

WHITE, M, WALKER, J, GLONEK, G, and BURNS, N (2000a), "Re-investigation of the Effectiveness of the Victorian TAC's Road Safety Campaigns". Paper submitted to Road Safety Research, Policing and Education Conference, Brisbane, November 2000.

WHITE, M, WALKER, J, GLONEK, G, and BURNS, N (2000b), "Re-investigation of the Effectiveness of the Victorian Transport Accident Commission's Road Safety Campaigns". Report No. 4/2000, Safety Strategy, Transport South Australia, November 2000.

RESPONSE BY MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE

TO

"Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns"
(White, Walker, Glonek and Burns, November 2000)

INTRODUCTION

This document is the response by Monash University Accident Research Centre (MUARC) to an investigation of some specific MUARC research. MUARC welcomes reviews of its work because of the critical importance of many of MUARC's results and conclusions. MUARC endeavours to conduct its research to the highest possible scientific standards, while recognising the need to provide results and advice in a timely manner. When necessary, MUARC always qualifies any of its work which is less than definitive.

The investigation relates principally to MUARC's research on traffic enforcement and road safety advertising in Victoria. MUARC is independent of any policy considerations which favour or disfavour various strategic approaches to the use of these measures to achieve road safety goals. MUARC's aim, within the constraint of the resources available to it, is to provide objective information on the effectiveness and cost-effectiveness of these measures in the broader context of factors affecting road trauma. Scientific principles have always been applied by MUARC to achieve this aim.

THE RE-INVESTIGATION PROCESS

In April 1997, the South Australian Office of Road Safety wrote to the Director of MUARC requesting that the data used for MUARC report no. 52 (RN52), *Evaluation of Transport Accident Commission road safety television advertising* (Cameron, Haworth, Oxley, Newstead and Le 1993), be supplied to the Office so that testing of MUARC's statistical models could be carried out by Adelaide University. In July 1997, the Transport Accident Commission (TAC) authorised the Office to obtain the advertising data from MUARC, subject to provisos including consultation with MUARC. In October 1997, the Office of Road Safety wrote to TAC accepting these provisos and in November 1997, MUARC provided this data to the Office. The Victoria Police authorised the provision of the enforcement data.

Since that time, MUARC has been provided with eight draft reports produced during the investigation by the Office of Road Safety, namely:

- *Re-analysis of the MUARC data on the efficacy of television advertising* (N.Burns), April 1998 (56 pages)
- *Re-investigation of Victorian data on the efficacy of TAC road safety advertising campaigns* (M. White, N. Burns, J. Walker), 20 July 1998 (177 pages)
- *Re-investigation of data on the efficacy of the Victorian Transport Accident Commission's road safety television advertising campaigns* (M. White, N. Burns, J. Walker), January 1999 (111 pages)

- Same title, same authors, March 1999 (102 pages)
- Same title, same authors, 24 March 1999 (103 pages)
- *Re-investigation of data on the effectiveness of the Victorian Transport Accident Commission's road safety television advertising campaigns* (M. White, J. Walker, G. Glonek, N. Burns), 26 April 2000 (101 pages)
- Same title, same authors, June 2000 (120 pages)
- *Re-investigation of data on the effectiveness of the Victorian Transport Accident Commission's road safety campaigns* (M. White, J. Walker, G. Glonek, N. Burns), November 2000 (119 pages)

The first two of these reports focused on the data used in MUARC report RN52. The third and subsequent reports focused on the data from MUARC report RN74, *Modelling of some major factors influencing road trauma trends in Victoria 1989-93* (Newstead, Cameron, Gantzer and Vulcan 1995).

A paper by White, Walker, Glonek and Burns, *Re-investigation of the Effectiveness of the Victorian TAC's Road Safety Campaigns,* has been published in the proceedings of the Road Safety Research, Policing and Education Conference, Brisbane, 26-28 November 2000 (White et al 2000a). One aim of that paper was to summarise the findings of the final report on the investigation, released by the Office of Road Safety in November 2000 (White et al 2000b).

This document aims to provide MUARC's response to both the paper and the report. Some of MUARC's response may relate to material or text which appears only in the report. MUARC's response will also refer to material in the earlier Office of Road Safety reports listed above, when there appears to be selective inattention to the inclusion of relevant material in White, Walker, Glonek and Burns' (November 2000) paper and final report.

White et al's (2000b) report includes an investigation of the analysis reported by Cameron and Vulcan (1998), *Evaluation review of the Supplementary Road Safety Package and its outcomes during the first two years,* in their role as private consultants to the Land Transport Safety Authority, New Zealand. This work was not associated with MUARC. Comments may be provided by the Land Transport Safety Authority and by the consultants in their private capacity.

MUARC RESEARCH ON ENFORCEMENT AND ADVERTISING

White et al (2000b) claim that "Although a number of different statistical techniques were employed by MUARC in their earlier evaluations [of enforcement and media campaigns], multiple linear regression became the favoured technique in their later work".

MUARC has completed thirteen substantial studies involving the analysis of real crash data in relation to Police enforcement, road safety advertising and sometimes additional factors. The MUARC reports on these studies, and the principal statistical framework used in each case, are listed in the table below.

No.	Authors	Report Title and Date	Statistical framework
RN29	Thoresen, Fry, Heiman & Cameron	*Linking economic activity, road safety countermeasures and other factors with the Victorian road toll* (July 1992)	multiple regression time series analysis
RN37	Drummond, Sullivan & Cavallo	*An evaluation of the random breath testing initiative in Victoria 1989-1990: Quasi-experimental time series approach* (September 1992)	quasi-experimental time series evaluation
RN38	Cameron, Cavallo & Sullivan	*Evaluation of the random breath testing initiative in Victoria: Multivariate time series approach* (September 1992)	quasi-experimental interrupted time series (with covariate) evaluation
RN42	Cameron, Cavallo & Gilbert	*Crash-based evaluation of the speed camera program in Victoria 1990-1991* (December 1992) *Phase 1: General effects* *Phase 2: Effects of program mechanisms*	(1) quasi-experimental interrupted time series (with covariate) evaluation (2) multiple regression time series analysis
RN52	Cameron, Haworth, Oxley, Newstead & Le	*Evaluation of Transport Accident Commission road safety television advertising* (September 1993) *Part 1: Overview and general effects* *Part 2: Economic analysis of the effects on crashes of the TAC advertising supporting enforcement* *Part 3:Effects of the Concentrate or Kill advertising campaign not directly linked to enforcement*	(1) exploratory trend analysis (2) multiple regression time series analysis (with stated assumptions) and economic evaluation (3) quasi-experimental before/after comparison evaluation
RN54	Rogerson, Newstead, & Cameron	*Evaluation of the speed camera program in Victoria 1990-1991* (February 1994) *Phase 3: Localised effects on casualty crashes and crash severity* *Phase 4: General effects on speed*	(3) quasi-experimental comparison evaluation (4) not based on real crash data
RN74	Newstead, Cameron, Gantzer & Vulcan	*Modelling of some major factors influencing road trauma trends in Victoria 1989-93* (July 1995)	multiple regression time series analysis
RN78	Newstead, Mullan & Cameron	*Evaluation of the speed camera program in Victoria 1990-1993: Phase 5: Further investigation of localised effects on casualty crash frequency and severity* (October 1995)	quasi-experimental comparison evaluation
RN126	Cameron, Diamantopoulou, Mullan, Dyte & Gantzer	*Evaluation of the country random breath testing and publicity program in Victoria 1993-1994* (October 1997)	quasi-experimental before/after comparison evaluation
RN129	Newstead, Cameron & Narayan	*Further modelling of some major factors influencing road trauma trends in Victoria: 1990-96* (April 1998)	multiple regression time series analysis
RN141	Diamantopoulou, Cameron & Shtifelman	*Evaluation of moving mode radar for speed enforcement in Victoria, 1995-1997* (October 1998)	quasi-experimental before/after comparison evaluation
RN149	Newstead, Cameron & Leggett	*Evaluation of the Queensland Random Road Watch program* (February 1999)	quasi-experimental interrupted time series evaluation
NA	Harrison, Fitzharris, Newstead, Gelb, Diamantopoulou & Cameron	*Evaluation of the effect of the deployment of hand-held laser speed-detection devices in the Melbourne metropolitan area* (September 1999)	quasi-experimental before/after comparison evaluation

In summary, five of the MUARC studies (or parts thereof) have used multiple regression time series analysis to link crashes with road safety program measures and other factors. The remaining ten studies and part-studies have all been quasi-experimental evaluations of the impacts of the programs, all but one of which were implemented in Victoria. The quasi-experimental time series evaluations were not critically dependent on the need to develop crash models including all influential factors in order to reach their conclusions.

The quasi-experimental evaluations were characterised by the inclusion of a control group of crashes to address the threat of "history" (ie. other causes operating at the same time). The time series based evaluations addressed the threat of "maturation" (ie. trends in the crashes over time). These are the two most critical threats to the integrity of evaluations of road crash countermeasures (Council et al 1980). Scientific evaluations are characterised by their active consideration of these two threats. In four of the MUARC evaluations the threat of "maturation" was not considered to be present, so a "before/after" comparison was used.

The above list of studies shows that multiple regression [time series analysis] has not been the favoured statistical technique used by MUARC in this area. Indeed, it has not been used in any study which MUARC has regarded as being a scientific evaluation of enforcement and/or advertising and has not needed to qualify the results. Multiple regression time series analysis has been used only in studies either attempting to represent the underlying mechanisms of road safety programs (eg. RN42, Phase 2, and RN52, Part 2) or studies attempting to consolidate previous findings (eg. RN74 and RN129). Each multiple regression analysis was based on prior scientific evaluations of the relevant road safety initiatives (or program of initiatives, such as enforcement and advertising aimed at drink-driving or at speeding).

White et al (2000a) claim that "Much of the MUARC research was brought together in a report by Newstead, Cameron, Gantzer and Vulcan (1995)", ie. RN74, and that this is "representative of the MUARC research that has been influential in shaping government policy on road safety television advertising levels" (White et al 2000b). As can be seen from the above list of MUARC studies, RN74 is not representative of the MUARC evaluation research on enforcement and advertising which either preceded it or followed it. As will also be discussed later, RN74 is not even representative of the research in RN52, which may have been influential in government policy.

MUARC macro-level trend analysis

MUARC has also conducted two studies of the trends in the Victorian annual road toll at the macroscopic level. These studies have indicated that the decrease in the Victorian road toll during 1990 and subsequent years was not consistent with expected trends and was a greater reduction than expected from pre-existing trends.

Smeed (1968) developed a relationship between fatalities per 10,000 vehicles and the vehicle ownership rate per 100,000 population after studying the time trends in many countries and also cross-sectional comparisons between countries. He found that the annual fatality rate was inversely related to the car ownership rate, with elasticity of about 2/3. RN74 (Appendix A) considered a modification to Smeed's "law" which allowed step changes in the level of the fatality rate during the year of introduction of key road safety initiatives in Victoria, with

Smeed's original relationship applying during the intervening years. Statistically significant step reductions were found during 1971, 1978, 1984 and 1990. The last of these years is the most relevant MUARC's response here because 1990 was the first full year in which the two new enforcement and advertising programs evaluated in RN38 and RN42 could be expected to have had a substantial effect.

Oppe (1989, 1991a,b) found that in many countries the fatality rate per million vehicle kilometres decreases as a negative exponential function with time. The total road travel usually increases according to a logistic saturation function with time, with perturbations reflecting economic conditions. Graham and Taylor (1994) found that the negative exponential function represented the NSW fatality rate during 1978-1993 very well. For Victoria, Cameron (1997) considered a modification to Oppe's negative exponential function which allowed step changes in level of the fatality rate during the years of key initiatives. Over the period 1979-1996, he found a statistically significant step reduction in the fatality rate during 1990.

Both studies have indicated that the general trend in the risk of road fatalities in Victoria had consistently improved during recent decades, as expected from increasing car ownership (Smeed) or exponential learning (Oppe), but commencing in 1990 there was a substantial reduction in this risk which could not be explained by pre-existing trends. MUARC does not consider these findings to constitute an evaluation of the 1990 initiatives, but the findings do provide supportive evidence of a real change in road safety in Victoria during the 1990's.

MUARC CONSULTANCY ADVICE

MUARC has been commissioned to provide road safety policy advice to jurisdictions outside Victoria. On occasions, this advice has been based on MUARC's evaluations of Victorian programs and/or the expert opinion of the nominated consultants. Reports from MUARC consultancy involvements which have been at least partly based on MUARC's evaluations of Victorian enforcement and advertising programs include:

- *Advice to assist bid for additional funding for road safety in New Zealand* (October 1994)
- *KwaZulu-Natal road traffic safety project: Assessment study* (August 1996)
- *Possibility of adapting some road safety measures successfully applied in Victoria to South Australia* (September 1996), RN102
- *Development of a five year action plan for community education on road safety in Western Australia* (November 1997)

In 1996, MUARC was commissioned by the South Australian Office of Road Safety to prepare the third report, RN102 (Vulcan, Cameron, Mullan and Dyte 1996). The proposal to the Office, included in RN 102, offered to "suggest additions or changes … which are likely to be effective" and nominated Professor Peter Vulcan and Max Cameron as advisors.

In their Executive Summary, White et al (2000b) state "The MUARC consultants … insisted that the review be extended to include road safety television advertising". This statement is not correct. As a consultant to the Office of Road Safety, MUARC was not in a position to insist on any aspect of the work. MUARC's proposal to the Office, which is included as Appendix D in RN102, suggested that the publicity supporting the South Australian

enforcement programs will be examined. It did not "insist" that this aspect be included. The Office of Road Safety could have declined MUARC's proposal. By accepting it, the Office implicitly accepted the suggestion that the supporting publicity be examined.

The study found that South Australia had a history of relatively low profile road safety advertising using non-emotive themes. The MUARC report recommended, *inter alia*, "that resources be allocated to double the exposure of television advertisements which support the speed camera and random breath testing programs".

Referring to this recommendation, White et al (2000a,b) state that "The main aim of the project reported here [the re-investigation] is to check the soundness of the advice concerning the effectiveness of high levels of television advertising, through re-analyses of the data originally analysed by Newstead et al (1995)".

White et al (2000a,b) do not mention that the consultancy report, RN102, makes no reference to Newstead et al (1995), RN74. Reference is made to RN52, but this was not the only basis on which the consultants advised their recommendation.

RN102 included a summary of the RN38 and RN42 evaluations which indicated the substantial reductions in severe crashes associated with the Victorian enforcement and advertising programs aimed at drink-driving and speeding. RN102 also noted that the level of road safety television advertising in Adelaide during 1993-1995 was less than half the level in Melbourne during the same years. Since South Australia had random breath testing and speed cameras, it was reasonable that similar effects as observed in Victoria could be achieved if South Australia supported these enforcement operations with road safety television advertising at the same level as Victoria, ie. approximately double current levels at the time. (RN102 also recommended an increase in random breath testing to equate with Victorian levels.)

Thus the consultants' advice in RN102 was not dependent on the findings in RN52. Indeed, if the findings in RN52 regarding advertising levels which are economically justified had been the basis of the consultants' recommendation, the recommendation would have been to more than double the level of advertising in South Australia. Finally, it should be noted that RN102 did not recommend that any increase in advertising should be based specifically on TAC-style advertisements; this issue was suggested as the basis of research in South Australia.

In all of MUARC's consultancy advice to other jurisdictions, including that related to road safety advertising in support of enforcement, reference has been made to the evaluation studies RN38 and RN42, as well as to RN52. This has been considered an appropriate background to the advice offered, given that RN52 includes a number of assumptions and qualifications on which it was based. RN38 and RN42 provided the scientific evidence of the effectiveness of the Victorian programs. RN52 provided an attempt to separate the contributions of the enforcement and advertising components, and on the basis that this had been achieved (with assumptions), provided an economic analysis of the advertising levels.

IMPORTANT DIFFERENCES BETWEEN RN52 AND RN74

White et al (2000b) state "It is considered that two selected RN74 analyses are representative of the MUARC research that has been influential in shaping government policy on road

safety television advertising levels". It is possible that RN52 may have been influential, if its stated assumptions were accepted by its readers, but as described in the previous section, all MUARC's consultancy advice in this area has been based on RN38 and RN42 as well. However, it is not correct that the research in RN74 is representative of that in RN52. There were important differences between these two research studies in their objectives, data analysed, and assumptions (see details below).

Report RN52

In early 1993, MUARC was commissioned by the Transport Accident Commission (TAC) to undertake an evaluation of the TAC's road safety television advertising. The TAC was facing decisions regarding the level of investment in road safety advertising it should make in the future. To assist these decisions, TAC sought advice on the level of effectiveness of the campaigns conducted to date; in particular the relationship between the level of advertising investment and reductions in road injury costs. The urgency of the study required that it be carried out over a 2½month period and that it would involve a number of assumptions.

The analysis reported in RN52 is in three parts, using the statistical techniques shown:

- *Part 1: General effects*: exploratory trend analysis
- *Part 2: Advertising supporting enforcement*: multiple regression time series analysis (with stated assumptions) and economic evaluation
- *Part 3: Concentrate or Kill advertising*: quasi-experimental before/after comparison evaluation

The results in Part 1 were inconclusive and Part 3 did not find evidence that the "Concentrate or Kill" advertisements reduced the crashes of their target group. Part 2 built on the evaluations reported in RN38 and RN42. Part 1 had indicated that it was not feasible to undertake a conclusive evaluation of the effects of the enforcement-related advertising alone.

It was assumed from RN38 and RN42, respectively, that the program involving increased random breath testing using "booze-buses" and supporting TAC advertising, and the program involving additional speed cameras and supporting TAC advertising, were each effective in reducing crashes. These were reasonable assumptions, given the scientific evaluation techniques used in these two studies. The methods, while complex, had precedents in the road trauma countermeasure evaluation literature (Garber and Graham 1990, Wagenaar, Streff and Schulz 1990). RN38 and RN42 have since been described as models of their type in the field (J. Broughton, TRL, UK; personal communication). These two studies will be outlined in greater detail later.

It was also assumed that the functional form of the mathematical model linking monthly casualty crashes with measures of the monthly operations of the speed camera program and the supporting TAC advertising, as found in Phase 2 of RN42, was correct. This assumption was more tenuous, but had its basis in a review by Hakim et al (1991) which suggested that multiplicative (rather than additive) function models best represent road trauma series. In the context, the functional form used was consistent with the expectation of diminishing returns from increasing levels of enforcement and advertising, in terms of crash reductions achieved. Such functional forms linking crash reductions with road safety program inputs have been frequently used in New Zealand resource allocation models (Land Transport Safety Authority 1996, 1998a,b,c). Nevertheless, this assumption was clearly stated in RN52.

The final modelling results of Part 2 of RN52 were for monthly casualty crashes during "high alcohol hours" (HAH) and "low alcohol hours" (LAH) separately. The LAH model linked monthly crashes in Melbourne during 1983-91 with unemployment rates, speed camera infringements, speed camera hours, and speed-related TAC television advertising TARPs (Target Audience Rating Points – a measure of audience reach). This was an extension of the model found in Phase 2 of RN42. The HAH models linked monthly crashes in Melbourne and country Victoria during 1983-92 with unemployment rates, random breath tests (car- or bus-based tests), alcohol sales, and the drink driving-related television advertising Adstock (a measure of current and past TARPs, considered to represent current awareness). The HAH models were developed at a later stage of the project than the LAH models, after MUARC had become aware of the Adstock concept (Broadbent 1979).

The economic analyses to determine the points of diminishing returns of levels of speed-related and drink-driving television advertising were based on the advertising components of each of the models described above. The other factors (unemployment rates, alcohol sales and enforcement levels) were set at average levels. The key parameter used in the economic analysis for each type of advertising was the estimated exponent of the advertising variable (TARPs or Adstock) in the multiplicative model. White et al (2000a,b) have not questioned the economic analysis from this basis, and hence the economic analysis will not be discussed further here.

For each economic analysis, RN52 displayed the functional form of the relationship between advertising levels and casualty crashes (when all other explanatory factors in the model are held constant), but preceded this by the statement that the form of the model was assumed to be correct and that the analysis had been confined to estimating the exponent [elasticity] of each explanatory variable. It should be noted that neither RN52 nor any other MUARC publication uses the expression "dose-response relationships" (White et al 2000a,b) to describe these relationships. MUARC considers that this expression gives a misleading impression of the process through which the relationships shown in RN52 were developed.

The economic analysis was followed by a statement that the results were based on the following assumptions:

"(a) The form of the relationships fitted to monthly numbers of casualty crashes occurring in Victoria during the "low" and "high" alcohol hours, respectively, in 1983-92 was correct.

"(b) The explanatory variables included (where appropriate, considering the time of week of the crashes being modelled) in the fitted relationships in addition to levels of TAC road safety television publicity (measured by TARPs and Adstock), namely monthly levels of speed camera operating hours, speed camera TINs issued, random breath tests, alcohol sales, unemployment rates, trend and seasonality variables, did not omit any other major variables with substantial effects on casualty crashes during the corresponding periods of the week.

"(c) The correlations between monthly levels of TAC television publicity and other variables included in the fitted relationships did not have a major effect on the estimates of the publicity impacts used in the economic analyses.

"(d) The decay in awareness of the road safety messages from the television advertising follows a negative exponential function with a half-life of five weeks.

"(e) The cumulative awareness of current and previous advertising, considered to be measured by Adstock, did not reach a level where satiation of the messages or "wear-out" occurred in any month during 1989-92, resulting in a relatively low level of effectiveness of the advertisements compared with expected."

Hence, while in the economic analysis the MUARC researchers made use of the estimated elasticities (and the implicit functional form assumed for the models and hence for the relationship between advertising levels and crashes), RN52 clearly stated the assumptions on which the economic analysis requested by TAC was based. The reader was cautioned that assumptions had been made in a total of six places in the report, including the Executive Summary.

Notwithstanding the assumptions stated, the models in RN52 have been tested and the diagnostics have been found to be satisfactory. During the project described in RN52, there was limited diagnostic testing in the time available, but the models did explain at least 69-89% of the variation in monthly crashes and there was no evidence of serial correlation as measured by the Durbin-Watson test. Time did not permit full consideration of the implications of the inter-correlations between the explanatory variables in the models, and this was why assumption (c) was stated. White et al (2000b, section 5.3) have suggested that high collinearity between the advertising and enforcement variables used in RN74 is problematic, but they do not pursue the issue in that context.

However the issue of collinearity between the variables used in RN52 has been investigated by White et al and found not to be a problem. The Executive Summary of the July 1998 re-investigation report, which focused on the data and analysis methods used in RN52, stated:
"After conducting the necessary diagnostic tests, it was concluded that there was no serious violation of the underlying assumptions [of the models]. *Nor was there a problem of 'multicollinearity' (ie. high levels of inter-correlation among the predictor variables."*

After examining the correlations between each pair of the variables used in RN52, and finding that none exceeded 0.9, the July 1998 re-investigation report stated:
"If the strongly correlating variables are logically (ie. conceptually) distinct, as they are in this case, they can all be retained in the multiple regression analysis. The statistical problems do not arise until the correlations exceed about 0.90, which is not the case here".
Later the report described a formal test of the multicollinearity between the set of predictor variables included in each model in RN52. It was found that the test was acceptable in each case.

Thus, while the original testing was limited, the diagnostic testing by the authors of RN52 and by White et al themselves has together found that the models linking monthly casualty crashes with other variables were satisfactory. This diagnostic research suggests that the original assumptions in RN52 were sound.

Report RN74

As noted earlier, there were important differences between RN74, which has been the focus of White et al's later re-investigation reports, and RN52. These differences are such that

RN74 cannot be considered representative of RN52, let alone representative of MUARC's evaluation research on enforcement and advertising.

The aim of RN74 was not to evaluate the enforcement and advertising programs, in the sense of providing a scientifically-defensible estimate of the impact of the Victorian programs on crashes. That role had already been played by RN38 and RN42. Part 1 of RN52 had also indicated that it was not feasible to provide a conclusive evaluation of the separate effects of the enforcement and advertising initiatives. (Part 2 had attempted to separate the effects, but as discussed above, this process involved many assumptions.) RN74 included summaries of these evaluations as background, and provided references to the relevant reports.

The aim of RN74 was to combine the results of evaluations of five Victorian road safety programs (summarised by Vulcan 1993) and information on the influence of two socio-economic factors (unemployment rates and alcohol sales) to assess the contribution of each program/factor to Victorian road trauma trends during the early 1990's. To do this, the study had to rationalise the criterion crashes of the key evaluation studies (RN38, RN 42 and evaluations of Accident Black Spot treatments) to a common base, namely serious casualty crashes. These crashes represent a sub-set (about one-third) of the total casualty crashes which were the focus of the final models in RN52.

The narrower focus of the crashes considered for the analysis in RN74, compared with RN52, is the most critical difference between these two studies. Comments made on MUARC's analysis of serious casualty crash data in RN74 do not directly relate to the results from the casualty crash analysis in RN52. White et al (2000b) state that "It is considered unlikely that the main points to be made in this report would be different if any if any other MUARC analyses had been selected for re-investigation". MUARC's response is that it is not unlikely, if RN52 had been the focus, contrary to White et al's (2000b) opinion.

It should be noted that the time trends in serious casualty crashes in Victoria were not the same as those in casualty crashes during 1983-92, in particular the point where each monthly series appeared to peak during the period (see later discussion).

As a further step in rationalising the previous studies, RN74 analysed the serious casualty crashes, for both the LAH and HAH times of the week, over the years 1983-92 (initially), whereas RN52 had analysed the LAH casualty crashes over the years 1983-91. Also, RN74 measured the levels of speed-related advertising in Adstock units, whereas RN52 measured this in TARPs. Finally, the RN 74 model for HAH crashes in Melbourne included the random breath tests conducted from buses only, whereas RN52 had considered the total monthly tests conducted from car-based and bus-based operations by the Victoria Police.

To achieve the aim of RN74, it was necessary to develop separate models of the monthly crashes in Melbourne and country Victoria, during HAH and LAH times, as functions of available data representing variations in the key elements of the road safety programs, and data representing influential socio-economic factors. This process was entirely overt in RN74. There was no suggestion that a scientific evaluation of each program element was being conducted. There was no use of a control group of crashes in an attempt to take into account other, unmeasured factors. There was not even a suggestion that all influential factors had been included in the models; the title of RN74 refers to "… some major factors …".

The results of RN74 are a direct extension of these models, which in turn were based on their assumed functional forms and the factors which were assumed to contribute in each case. Testing of the models was carried out in RN74 by extrapolating the initial models developed for the 1983-92 period to predict the serious casualty crashes during each month of 1993 and then comparing the predictions with the actual crashes. The prediction was found to be reasonable, so the models were re-estimated for the 1983-93 period. The revised models were used calculate the final results of RN74. It should be emphasised that the ability to predict was a demanding test of the models; the aim of RN74 was to *explain* the road trauma trends in Victoria in terms of the contributions of some factors, not to *predict* those trends. However, the satisfactory prediction suggests that the models included most of the key factors.

Essentially the same models of serious casualty crashes during the 1983-93 period were tested for serial correlation and heteroskedasticity of the residuals by Fry (1996). He concluded that the models fit well and appear well specified. It is noteworthy that Fry's paper was published in the internationally-recognised journal *Accident Analysis and Prevention*, which requires technical peer-review of all its papers before acceptance. The modelling methods used in RN74 were published by Cameron, Newstead and Vulcan (1994) in *Proceedings of the 17th Australian Road Research Board Conference*, which also requires technical peer-review of its papers before acceptance.

Apart from suggesting that high collinearity between the advertising and enforcement variables used in RN74 is problematic, White et al (2000b) do not report any diagnostic testing of collinearity problems along the lines of the testing they carried out on the data used in RN52. It is unclear whether they have not carried out such testing or whether the testing failed to identify any modelling problems associated with collinearity. This omission from the re-investigation of the RN74 data is unfortunate. The results may have contributed to the above findings which suggest that the diagnostic testing of the RN74 models was satisfactory.

As a final remark, it should be noted that RN74 did not have a specific focus on the advertising component of the Victorian road safety program, unlike RN52. While RN74 did estimate elasticities between the advertising variables and crashes, it did not suggest or display any relationships based on these elasticities. Nor did it undertake any economic analysis of the points of diminishing returns of advertising levels. This appears only in RN52.

Summary of differences

There are important differences between RN52 and RN74 in their objectives, the type of crashes analysed, the time periods covered, the treatment of levels of speed-related advertising, and the inclusion of car-based random breath testing. The economic analysis in RN52 was based on models of all casualty crashes (not just serious casualty crashes), the period 1983-1991 for the LAH casualty crashes (not 1983-1992; this period was used for the HAH crashes), the levels of speed-related advertising were measured in TARPs (not Adstock; this was used only for drink-driving advertising), and the random breath test data included car-based tests (not just the bus-based tests included in the model in RN74).

These differences invalidate White et al's (2000b) assertion that the RN74 analysis is representative of that in RN52.

DATA DREDGING

White et al (2000a) claim that "The methods employed for the selection of variables in the MUARC modelling process can be described as 'data dredging' ". To support this claim, White et al (2000b) quote principally from RN29, *Linking economic activity, road safety countermeasures and other factors with the Victorian road toll* (July 1992).

Report RN29

RN29 was carried out by ARRB Transport Research and the Monash University Department of Econometrics in conjunction with MUARC. Its initial aim was to investigate the relationship between the level of economic activity and the road death toll. The specific objective was to link the Victorian road toll with economic, road safety, social and other factors using historical data to identify the relevance of the various factors. A literature review conducted prior to the research indicated that statistical models of the road toll should be based on an appropriate conceptual model with the ability to describe, explain and predict the phenomenon being analysed and to incorporate policy variables. The literature review also showed that in such models typically only one variable is included to describe the economic conditions and this is usually chosen to be some measure of unemployment.

The "road trauma chain" developed by MUARC was used as the conceptual model in RN29. This "chain" was used to categorise the 36 potential explanatory variables on which data had been collected into groups of variables which were considered to be related to specific steps in the chain. The aim was to consider competing explanatory variables only from those in each group and to chose one variable to represent the group in the statistical models. Because it was unclear whether explanatory variables may be linked to the road toll in the current month or influence that in the next month, each variable in a group was also considered in one-month lagged form.

White et al's (2000b) selective quotation from RN29 gives the impression that 66 potential explanatory variables, each in both raw and lagged form, were considered in RN29 without any conceptual structure to the choice of each for inclusion. The authors of RN29 were well aware of the dangers of potential spurious inclusion of variables in the statistical models if such a large number of candidate variables had been allowed to compete among themselves in a standard multiple regression analysis. It was for this very reason that the variables were considered within the context of "road trauma chain" in order to minimise this danger.

Nevertheless, it is possible that the explanatory variables ultimately chosen may have included some which were not optimal in the sense that were not the closest to the underlying factor explaining the variations in each variable in a group, but were selected by the multiple regression for partly chance reasons. To some extent, MUARC agrees with White et al (2000a) that the models developed in RN29 "should be considered to be 'hypothesis generating' rather than 'hypothesis testing' ". It is for this reason that MUARC does not consider RN29 to be a scientific evaluation of the impact of the countermeasures included in the RN29 models of Victorian road fatalities. However, the hypotheses "generated" from RN29 about likely influential factors on road trauma generally in Victoria, especially the contribution of economic activity measured by unemployment, were instrumental in developing the methods for testing the hypotheses (for presence or absence of effect of each program) in the scientific evaluations described in RN38 and RN42 (see below).

It should be noted that RN29 analysed only numbers of fatalities (in total and in three road user groups) and did not include any measure of TAC road safety advertising. For these reasons alone, it could not be considered to have tested any hypotheses related to factors contributing to the reduction in casualty crashes (or even serious casualty crashes, since fatal crashes are a small proportion of these) in Victoria during the 1990's, especially the contribution of the advertising. The fatality models were developed for crashes only to December 1990, but were tested (reasonably successfully) on the 1991 monthly road tolls.

Reports RN37 and RN38

RN37 was the first of MUARC's quasi-experimental evaluations of the "booze bus" initiative and supporting TAC publicity introduced in late 1989. This study made use of Sydney to provide a control group of crashes for Melbourne, and a relatively unaffected rural region of Victoria to provide the control crashes for the rural region where buses had operated during 1990. In each of these four areas, HAH serious casualty crashes up to the start of the initiative were modelled using univariate ARIMA time series methods. The univariate models were then projected to estimate the expected crashes during the post-implementation period. The actual and expected crashes were then compared, and the differences in the areas influenced by the program, compared with the uninfluenced areas, were used to make inferences about the effect of the program on the crashes up to the end of 1990.

RN38 was also a quasi-experimental evaluation of the effect of the "booze bus" initiative and supporting publicity, focusing on HAH serious casualty crashes to the end of 1991. Based on a recently developed method of estimating vehicle travel from fuel sales (Lambert 1992), it was noted that the rate of increase of travel had levelled out and then fallen in Victoria during 1990-1991, whereas it had remained steady in NSW. It also noted that unemployment rates had begun to increase substantially in Melbourne and rural Victoria, respectively, during 1989-1990, but parallel increases in NSW did not commence until about a year later. These differences between Victoria and NSW cast doubt on the integrity of the control groups of crashes from Sydney and rural NSW to address the threat of "history" (other causes operating at the same time) when assessing the effects of the program in Melbourne and rural Victoria. (It had been decided that, since the "booze bus" operations had expanded throughout rural Victoria during 1991, rural Victoria would be considered as a whole in RN38.)

To overcome this methodological problem, it was decided to take these differences between the two States into account explicitly by including a measure of vehicle travel and/or unemployment rates in the evaluation time series models as "covariates". It was noted that the two potential covariates were correlated. RN38 chose to use unemployment rates as the preferred covariate, in part because separate rates were available for both the metropolitan and rural areas of Victoria and NSW. There are also theories that unemployment rate is related to the quality of road use, in that during times of high unemployment, travel at night and by young drivers, both of which tend to be associated with relatively high crash risks, tend to be reduced to a relatively greater extent (Harry 1997).

Multivariate time series models of monthly HAH serious casualty crashes in each of the four areas were developed. The models were considered multivariate because, in addition to time-related factors (trend and seasonality), they included the covariate (unemployment rate in the area) as an explanatory factor for the crash frequencies. The multivariate models also included explanatory factors in the form of step functions reflecting hypothesised changes in

the crash level during three post-implementation periods (December 1989; 1990; and 1991). Based on this framework, the multivariate models were developed using two methods: ARIMA Intervention Analysis (Box and Tiao 1975), and multiple regression analysis assuming a multiplicative structural form. These methods were both "interrupted" time series analyses because the step functions had the potential to interrupt the level of the crash series. It should be emphasised that all of the time series models were developed in mirror image form, no matter whether the focus was on crashes in each area of Victoria (where the program effects were hypothesised) or in NSW. The inferences regarding the effects of the program in Victoria were based on the final step in which the estimated step changes in Victoria were compared with estimated step changes in NSW.

Report RN42

Phase 1 of RN42 was a quasi-experimental evaluation of the increased number of speed cameras, which were announced in December 1989, and the supporting TAC speed-related advertising, which was launched in April 1990. Its approach was very similar to RN38, except that it focused on LAH casualty crashes and, separately, their severity defined as the proportion involving death or serious injury. The LAH crashes were chosen as the focus because a subset of crashes was sought which was relatively immune from the effects of the RBT "booze bus" program which had commenced about the same time. LAH crashes were not considered to identify all speeding-related crashes. Speeding was considered likely to be involved in a large proportion of HAH crashes, as was drink-driving, but it was not considered feasible to scientifically evaluate the general effect of the speed camera program on HAH crashes separately from the effect of the "booze bus" program on those crashes. (However, when the localised effects of the speed camera program were evaluated in RN54 and RN78, they were found among both LAH and HAH crashes. This later finding did not invalidate the conclusions of RN42 with respect to the general effects of the program on LAH crashes. RN42 acknowledged the necessary limitation in its scope.)

As in RN38, the multivariate time series models of monthly LAH casualty crash frequency and severity included unemployment rate as a covariate. The models also included step functions representing hypothesised changes in level of each series at each key stage during the implementation of the Victorian program. The rationale for preferring unemployment rate instead of vehicle travel as the covariate in the LAH crash models was not based on any theory linking economic conditions with discretionary, perhaps higher risk, travel which may have been a more frequent part of HAH travel. However, it was also not based on consideration of greater association between unemployment rates and LAH crashes, compared with that between vehicle travel and LAH crashes. Nevertheless, it was found that unemployment rate was often a statistically significant factor in each of the models of LAH crash frequency and severity.

Phase 2 of RN42 had the objective of attempting to understand the underlying mechanisms of the speed camera program (camera operations and supporting TAC publicity) which, so far as its general effects were concerned, was considered to have been scientifically evaluated in Phase 1. It used multiple regression time series analysis to develop models of Melbourne casualty crash frequency and severity as functions of unemployment rates, road safety advertising TARPs, speed camera hours, and camera-related speeding tickets issued during the month. This analysis formed the basis of the analysis in Part 2 of RN52, and has been discussed earlier. During Phase 2 of RN42 there was no consideration of alternative explanatory factors (for which data was available) for inclusion in the models.

Report RN52

Part 2 of RN52 built on Phase 2 of RN42 and did not consider any alternative explanatory factors for LAH casualty crashes in Melbourne, apart from the possibility that the TARPs of TAC advertisements with the "concentration" theme may add explanatory value to those with a pure speed-related theme. This possibility arose because the principal "concentration" advertisement used in the early 1990s, *Country Kids*, included speeding images and references to speed cameras, and may have been confused and interpreted as a speed-related message. (A broader effect of this advertisement was considered possible, notwithstanding that Part 3 of RN52 did not find evidence that the concentration advertisements reduced the crashes of their specific target group.)

Regarding the models for HAH casualty crashes in Part 2 of RN52, this research was carried out after MUARC became aware of the Adstock method of representing current and past TARPs as a single value in each period. This had been found to be better correlated with current awareness of a television advertisement than raw TARPs in the current period (Broadbent 1979). MUARC considered that it was likely that Adstock would be more closely associated with relevant crashes than the TARPs in a period (if there was any relationship at all) and hence Adstock became MUARC's preferred measure of television advertising levels, notwithstanding that it required assumptions about the retained awareness half-life (since investigated in RN126 and Shtifelman, Cameron and Diamantopoulou 1999). The final HAH crash models on which the economic analyses were based were those which considered Adstock as the measure of advertising. The TARPs models were provided for comparison with the LAH casualty crash models which had been developed considering only the TARPs measure of advertising. The project time frame did not allow consideration of an Adstock measure for speed-related advertising.

Summary of MUARC processes in variable selection

MUARC's initial analysis of factors linked to road trauma trends in Victoria examined relationships with monthly fatalities only and considered the potential explanatory variables within a conceptual framework used to minimise any spurious selection due to chance. RN29 was not considered to be a scientific evaluation of the factors. It did suggest important influential factors, such as those representing economic conditions, which needed to be taken into account in MUARC's subsequent evaluations of Victoria's road safety initiatives.

The quasi-experimental time series evaluations of the RBT "booze bus" initiative and the increased speed cameras, with the supporting publicity in each case, needed to take into account the different trends in vehicle travel and/or unemployment rates in Victoria and NSW to ensure the integrity of the evaluation design. Unemployment rate was chosen for non-arbitrary reasons associated with the specificity of the data and theories about a causal role.

The multiple regression time series analyses continued to use unemployment rates as an explanatory factor without consideration of alternative economic variables. Initially the models used TARPs to represent road safety television advertising as an explanatory factor, but when MUARC became aware of Adstock and its conceptually better basis for a link with awareness and hence potentially with crashes, Adstock became the preferred measure.

This process of evolution of the choice of covariates and explanatory factors for inclusion in MUARC's quasi-experimental evaluations and crash modelling analyses, respectively, did not constitute "data dredging" as implied by White et al (2000a,b). The factors were chosen with careful attention to avoiding spurious inclusion and on the basis of reasoned consideration of their possible causality and potential explanatory role.

Professor Eccleston's review

White et al (2000b) quote a review of their re-investigation reports by Professor John Eccleston, Professor of Statistics and Probability, Department of Mathematics, University of Queensland. White et al state "The methods employed for the selection of variables in the RN74 modelling process can be described as 'data dredging' (Eccleston, 2000)".

Reference to Eccleston's (2000) review indicates that his brief was "to review the appropriateness of the statistical arguments and procedures used in the Transport SA reports". He did not have a brief to review the relevant MUARC reports and there is no reference to him being provided with or obtaining these reports. His review document states "The methods employed for the selection of variables in the RN74 modelling process, as reported in subsection 7.3 [of White et al 2000b], is data mining or data dredging" (underlining added).

MUARC has concerns about the extent of the information about MUARC's research available to Professor Eccleston as the basis for this statement. As outlined above, White et al (2000b) have not described MUARC's processes fully nor accurately. MUARC questions whether an opportunity to gain a full understanding of MUARC's research program in the area was provided to Professor Eccelston or other reviewers of White et al's reports.

"Data dredging" and the White et al adjustment for economic activity

White et al (2000b) have proposed an alternative to unemployment rate as a measure of economic activity and then conclude "that the strong countermeasure effects reported by Newstead et al [RN74] are largely a result of their arbitrary decision to use the Unemployment variable as their measure of economic activity". White et al's current alternative measure of economic activity, and the implications for conclusions, will be discussed later.

It is worth noting that White et al have apparently considered a range of indicators for this role, with a range of lags, during their investigation. In their 1999 re-investigation reports they discuss various indicators of economic activity and then propose that the monthly unemployment rate, brought forward by 12 months, would be adequate in providing the sole explanation for the trends in both the LAH and HAH serious casualty crashes during 1983-92.

When discussing the economic indicators, White et al considered unemployment rate to be a *coincident* indicator, ie. it responds at much the same time as the general economy. Unemployment rates in Melbourne appeared to start rising during 1989-1990, and total vehicle travel in Victoria appeared to level out and commence falling in the same period. This suggests that unemployment rate is coincident with road use levels.

On this basis, MUARC responded in April 1999 that it is not justified to consider unemployment rate as a 12 month lagging indicator for road crashes, as White et al had done. It appeared to MUARC that White et al had considered the time trends in the road crash series and had attempted to match its pattern with that of the other available series, finding a match with the unemployment rate series when it was brought forward by 12 months.

MUARC does not consider it is appropriate to use "data dredging" to select a factor to explain the variation in a road crash data series. If sufficient factors, with a variety of leads and lags, are considered the analyst is almost certain to find a factor which explains the series well. However this approach runs the danger that the selected factor has no causal basis for the explanation, and that the apparent relationship is spurious. The use of a conceptual or theoretical model to select the factors for consideration is important to minimise this danger.

WHITE ET AL'S ALTERNATIVE INDEX OF ECONOMIC ACTIVITY

White et al (2000a,b) have proposed the Leading Index (LI) of Economic Indicators for inclusion in the models in RN74 instead of unemployment rate. While not accepting the relevance of RN74 to any other MUARC research on enforcement and advertising apart from itself (and perhaps RN129), MUARC questions the process through which White et al have chosen and used this variable as well as the conclusions they reach.

White et al (2000b) present arguments for the choice of LI which include "The peak in the Leading Index occurs at much the same time as the peak in all (ie, LAH + HAH) casualty crashes [sic? serious casualty crashes]. Therefore, it seems possible that the Leading Index could provide the best-fitting economic variable to use in statistical modelling of crash numbers". MUARC asks: What would White et al have proposed had LI not peaked at the same time as the crashes? Were other economic indicators considered that did not peak at this time? Was the aim to find an indicator which mirrored the crashes as closely as possible?

MUARC also cautions that the inferences which are made from pair-wise comparisons of time series variables can be very misleading when the relationships between crashes and other factors are truly multivariate. The peak in the crash series may change after adjustment for the influence of other factors, and then may not correspond with the peak of the specific explanatory variable under consideration. This issue will be raised again later in the context of the turnaround in crashes in Victoria during the late 1980's.

In a later section when commenting on the LI variable, White et al (2000b) state "[A] reason for selecting this measure was, admittedly, because it peaked at much the same time as the peak in crash numbers. In that respect the measure is a product of 'data dredging', … "

White et al (2000b) also state "Because the month-to-month variation in the Leading Index was shown (in analyses not reported here) to be unrelated to the month-to-month variation in crash numbers, a smoother version of the Leading Index was created by taking the 12-month centred-moving-average (CMA 12)." These are remarkable statements about the choice and use of an economic indicator to replace the one which MUARC has found in at least six studies to have statistically significant associations with monthly road trauma levels in Victoria.

Although LI was chosen as one which peaked and followed the general trend of crashes in Victoria, it apparently is not related to crashes at a monthly level. For this reason, it was replaced by White et al by its 12-month centred-moving-average before consideration in the RN74 models. A centred-moving-average would include values of LI from one month up to six months ahead of each specific month in which the crashes and explanatory factors are potentially related through the statistical model. In this analysis situation, MUARC questions how future LI values (up to six months ahead) could be causally related to the crashes in a specific month? Is the relationship White et al have found purely correlational?

When White et al (2000b) included the CMA 12 Leading Index in the RN74 models, the magnitudes of the estimated coefficients of the enforcement and advertising variables, and their statistical significance levels, were reduced. MUARC considers this to be an inevitable outcome of a process whereby an alternative economic indicator has been chosen on the basis of its coinciding peak and general correlation with the crash series. In these circumstances, the new economic indicator could be expected to take up some of the explanatory role which the enforcement and advertising variables previously held. This could then be expected to reduce the estimates of the coefficients of those variables, as well as their significance levels.

It is noteworthy, however, that in White et al's (2000b) LAH serious casualty crash model which included the CMA 12 Leading Index (Table 8.3.1), both the speed-related enforcement variable (speed camera tickets issued) and the advertising variable (speed and concentration Adstock) were statistically significant at the 5% level ($p < 0.05$). In White et al's HAH serious casualty crash model (Table 8.3.3) neither the drink-driving enforcement nor advertising variables were statistically significant, however the estimated coefficient of the advertising variable reduced from –0.015 to –0.013, a relatively small change. These findings suggest that, especially in the case of speed-related enforcement and advertising, there is evidence of a link with crashes even after the explanation associated with White et al's CMA 12 Leading Index is taken into account.

For reasons of their own, White et al (2000b) also consider a model of all (LAH plus HAH) Melbourne serious casualty crashes following the RN74 approach, with the CMA 12 Leading Index and the speed-related enforcement and advertising variables as explanatory factors (Table 8.3.2). This is mentioned here because of its relevance to MUARC's comments in the next section. The speed-related advertising was found to be statistically significant ($p = 0.006$), whereas the enforcement variable was not. The proportion of the monthly variation in crashes explained by this model (R^2) was 88.1%.

WHITE ET AL'S THREE-FACTOR MODEL

White et al (2000a,b) have proposed that a three-factor model incorporating (1) linear trend, (2) seasonality, and (3) the CMA 12 Leading Index would be adequate to explain the monthly variations in all Melbourne serious casualty crashes during 1983-1992. They discount the contribution of the speed-related advertising to explaining variations in the same crash series, as reported in their Table 8.3.2, saying that "speed advertising made only a weak contribution". In fact, Table 8.3.2 shows that the contribution of the speed-related advertising was statistically significant ($p = 0.006$) and that the magnitude of the estimated coefficient was –0.016, which is slightly greater than that estimated for speed-related advertising in RN52 and on which its economic analysis was based.

Again for reasons of their own, White et al (2000b) have presented the results of fitting an additive model of their three factors to the monthly crashes, rather than a multiplicative model. The multiplicative model form has been used throughout MUARC's multiple regression time series analysis. (It is made additive, prior to the use of multiple linear regression for fitting, by logarithm transformations of the crashes and explanatory factors.) The additive model cannot be claimed to be more parsimonious than a multiplicative model because both could include the same explanatory factors; only the form would be different. An extensive review by Hakim et al (1991) suggested that multiplicative (rather than additive) function models best represent road trauma series.

White et al (2000b) appear to suggest that an additive model is simple, and that this is a desirable feature compared with the multiplicative models used by MUARC. However, it should be noted that, since the models aim to represent monthly counts of road crashes, by definition their functional form must not allow negative values. The multiplicative functional form used by MUARC ensures that the number of crashes predicted by the model is constrained to be non-negative. An additive model would not necessarily meet this constraint. This mis-specification may lead to incorrect conclusions about the significance of factors included in the model.

The additive model of the three factors explained 87.4% of the monthly variation in the Melbourne serious casualty crashes. This is less than the 88.1% variation explained by the multiplicative model shown in White et al's (2000b) Table 8.3.2 which incorporated the same three factors plus the statistically significant speed-related advertising variable. White et al (2000b) state that when the enforcement and advertising variables, both speed and drink-driving related, were considered for inclusion in their additive model, none of these factors were statistically significant.

MUARC questions the change in functional form of the crash model for this analysis. If the multiplicative form had been retained by White et al (2000b), their findings in Table 8.3.2 show that, had the speed-related advertising been considered for inclusion in the three-factor model, it would have been statistically significant. In other words, it would have increased the proportion of monthly variation in crashes explained by the model by an amount greater than could be explained by chance. Thus the three-factor model could not be described as parsimonious, because it does not take into account at least one factor (ie. speed-related advertising) known also to be associated with crashes. A change in functional form, resulting in the contribution of this factor no longer being statistically significant, does not make the additive model parsimonious.

TIMING OF TURN-AROUND(S) IN VICTORIAN CRASHES

White et al (2000a,b) state that "Reports from some Victorian and overseas road safety agencies give the clear impression that the agencies believed that the TAC-funded enforcement and advertising campaigns were largely, if not entirely, responsible for *halting the increase* in crash numbers and for *initiating their decline*."

White et al (2000a) do not state, but White et al (2000b) do, that these "misunderstandings" [White et al's words] "are in no way attributable to the work of the MUARC researchers". The most that MUARC has claimed (eg. RN74) is that a number of road safety measures and

other factors have *contributed* to the reductions in road trauma in Victoria during 1990 and later years.

White et al (2000a,b) claim that the start of the decline in serious casualty crashes preceded the launch of the speeding and drink-driving enforcement and advertising campaigns, and that this casts doubt on their causal role. They claim this after comparing the time series variations in the each of the four activity measures with the trend in the relevant crashes in each case.

This analysis may be valid if a single factor (ie. the relevant enforcement or advertising) had been hypothesised as responsible for the trends in the relevant crashes during 1983-92. However, in RN74, MUARC proposed multi-factor models for the trends in crashes during this period and did not claim that any one measure was the only influential factor. Nor did MUARC claim that any program of measures (eg. related enforcement and advertising) was the only influence on the relevant crashes. The MUARC models in RN52 and RN74 all included socio-economic factors as well as road safety measures. The single-factor comparisons made by White et al (2000a,b) cannot be used to judge the contribution of each road safety measure to these multi-factor models, because they ignore the simultaneous effects the other variables have had on the observed crash series.

RN52 shows that the models for casualty crashes appear to represent the trends in the data well. The peaks in these crashes appear to be in 1989 for both HAH and LAH, and the substantial drops commencing in 1990 are consistent with the introduction of the two major enforcement and publicity programs commencing late-1989/early-1990. Nevertheless, the statistical significance of unemployment rate, alcohol sales (HAH model only), trend and seasonality indicates that these other factors had substantial explanatory roles both before and after the two programs commenced.

The models for serious casualty crashes in RN74 also appear to represent the trends in the data well. However the peaks in these crashes appear to be in 1988 for both HAH and LAH. The explanation appears to lie in a greater effect (higher elasticity) of alcohol sales on HAH serious casualty crashes compared with its effect on all HAH casualty crashes, and a greater effect (higher elasticity) of speed camera operations (tickets issued) on LAH serious casualty crashes compared with its effect on all LAH casualty crashes. Alcohol sales during 1989 were lower than the previous five years and the start of an annual decline to at least 1992. Speed cameras had been introduced in small numbers in Victoria during 1985 and these were active in 1989 as well as in the early months of 1990 before the 54 new cameras started to be delivered. Both of these factors could be expected to have crash reduction effects which were greater for serious casualty crashes compared with the effects on all casualty crashes.

While MUARC offers this explanation for the apparent different trends in serious casualty crashes compared with all casualty crashes (of which serious casualty crashes were about one-third), it is emphasised that MUARC's research has been confined to assessing factors which contributed to the reductions in road trauma in Victoria during the 1990's. MUARC has not evaluated the factors which may have been responsible for the turnaround in crashes of each level of severity prior to 1990.

TESTS OF QUANTITATIVE RELATIONSHIPS

White et al's (2000) test

White et al (2000a,b) suggest that RN52 and RN74 claimed that quantitative relationships between TAC advertising levels and crashes had been identified. They present the results of tests based on the data in RN74 which they claim fail to provide evidence of any quantitative relationships. White et al (2000b) state that the tests were replicated (with the same results) on the data in RN52, which they also state to be the source of the original "dose-response" [their words] curves, however they do not present the test results. MUARC considers this omission to be unfortunate, given that, of the two reports, only RN52 is relevant to advice about levels of road safety advertising.

It should be noted that MUARC has never used the expression "dose-response" to describe the relationships displayed in RN52. MUARC considers that this expression gives a misleading impression of the process through which the relationships shown in RN52 were developed, which was fully described (together with the assumptions made) in RN52.

The tests on the RN74 data involved defining a dichotomous (0, 1) variable for each of the four speeding and drink-driving enforcement and advertising measures and then including each pair (dichotomous and raw variable) together in the original (RN74) multiple regression analysis of the relevant serious casualty crashes. The dichotomous variable represented the impact of the road safety initiative as if it produced a step change in crash frequency at the time of its implementation. In each regression analysis of LAH and HAH crashes, respectively, the formerly statistically significant enforcement and advertising variables (in raw form) became non-significant and the dichotomous variables were also non-significant in every case. White et al (2000a,b) argued that, if the relationships had been truly quantitative, the raw measures should have retained their statistical significance. Failure to do so constituted a lack of evidence of a quantitative relationship, in White et al's opinion.

MUARC does <u>not</u> agree that the test performed by White et al (2000a,b) is an adequate test for the presence or absence of any quantitative relationships. It is known that if two highly correlated variables are included together in a multiple regression, the standard errors of the estimated coefficients of each variable will be substantially higher than if the variables are considered alone, and hence the estimates will be less reliable. More importantly, the estimated coefficients of the highly correlated variables will be individually meaningless in their interpretation. Because of the way the dichotomous variables were created, MUARC expects that they would be highly correlated with the raw measure. Tables 10.1 and 10.2 in White et al (2000b) show that the standard errors of the raw measures increased substantially when the dichotomous variables were added to the multiple regression analysis. By itself the increase in the standard errors would have reduced the likelihood of statistical significance of the raw measures.

In their 1999 re-investigation reports, White et al described the same tests on the RN74 data. In these reports, the correlations between each pair of raw and dichotomised variables were given, as follows:

- speed enforcement (camera tickets issued) 0.838
- speed-related advertising (Adstock) 0.999
- drink-driving enforcement (bus-based tests) 0.997
- drink-driving advertising (Adstock) 0.993

In the same reports, White et al state "there is little point in even attempting such an analysis if the correlations between the dichotomous and quantitative versions of the same variable are so high that statistical problems of collinearity would necessarily be introduced into the regression analysis. This could be the case for correlations above 0.90 (Tabachnik and Fidell, p. 96) and would almost certainly be the case for correlations of 0.99 or above".

For this reason, the 1999 re-investigation reports provided only the results of a regression analysis of LAH serious casualty crashes in which the speed enforcement was included as both raw and dichotomised variables, and the speed-related advertising was included in its original form (Table 5.5.2). The raw speed enforcement variable remained statistically significant ($p < 0.0005$) as did the advertising variable ($p < 0.0005$). The dichotomised enforcement variable was not statistically significant. White et al concluded that "It would seem that the original enforcement variable is a legitimate quantitative variable".

MUARC is surprised that, although White et al (2000b) had clearly recognised the problem of high collinearity in the context of their test, they decided to proceed with it and report the results. MUARC rejects the notion that the tests in Tables 10.1 and 10.2 are a valid test of the quantitative relationships. When the only valid test, according to White et al's own standards, was performed, White et al (1999) found evidence that the speed enforcement variable (camera tickets issued) has a quantitative relationship with LAH serious casualty crashes. MUARC considers that the presence or absence of quantitative relationships linking crashes with speed-related advertising and with drink-driving enforcement and advertising has not been adequately tested by the results in Tables 10.1 and 10.2 of White et al (2000b).

White et al's 1999 analysis

The 1999 re-investigation reports include additional analyses which indicate the presence of quantitative effects of the speed-related advertising and perhaps also the drink-driving enforcement. In those reports, White et al drew a distinction between the macro- and micro-level effects of the advertising and the enforcement on crashes. They claimed that the apparent effects of these factors were entirely due to the macro effects associated with their introduction. However, they also presented some evidence of the existence of micro-level effects (as they define it) after the introduction of the two enforcement and advertising programs. The micro-level effects could be described as quantitative.

In the case of LAH serious casualty crashes, White et al considered the period from August 1990 to December 1992 (29 months). They found negative correlations with the monthly speed camera tickets ($p = 0.097$) and speed-related Adstock ($p = 0.038$), the latter being statistically significant and the former nearly so. In the case of the HAH serious casualty crashes, they considered the period from January 1990 to December 1992 (36 months). For reasons of their own, they subtracted a linear downward trend from the crashes when examining correlations with the enforcement and advertising. They found a nearly statistically significant negative correlation with monthly bus-based random breath tests ($p = 0.056$) before the trend was removed from the crashes, which weakened after it was removed ($p = 0.375$). They also found a negative correlation with monthly drink-driving Adstock after the trend in crashes was removed, but this was not statistically significant ($p = 0.244$).

White et al's analysis of micro-level effects over at most three years of data was necessarily weaker than the original MUARC modelling over ten years, so it is not surprising that the

statistical significance of the tests for micro-level effects were weaker. Nevertheless, all of the correlations of monthly crashes with the enforcement and advertising variables were (appropriately) negative in sign, and one (speed-related advertising) was statistically significant.

White et al (2000) dismiss this earlier analysis as not being appropriate. MUARC accepts that, with the constraint on the amount of monthly data which could be analysed in this way, the analysis was necessarily weak, ie. the probability of detecting micro-level effects, if they really existed, was low. Nevertheless, the analysis did find statistically significant evidence of micro-level effects of the speed-related advertising, hence this should be reported.

In summary, White et al's (2000a,b) test of the presence or absence of quantitative relationships was not considered valid because of high collinearity problems associated with three of the four enforcement and advertising measures under consideration. When an apparently valid test had been performed on the fourth measure (speed camera tickets issued) in an earlier report, White et al concluded that it had a legitimate quantitative relationship with LAH serious casualty crashes. An alternative, albeit weaker, analysis has also indicated that there is a quantitative relationship between speed-related advertising and crashes.

STARTING MONTH OF DRINK-DRIVING ADVERTISING

White et al (2000b) have criticised MUARC researchers for using November 1989 as the start date of drink-driving advertising for the analysis behind RN52 and RN74, arguing that all other documents have used mid-December 1989 as the start of TAC road safety advertising.

As part of the Victorian Government's new road safety strategy launched in September 1989, responsibility for all road safety advertising in the mass media was transferred from VicRoads to the TAC (Vulcan and Cameron 1996). This had occurred by November 1989 ('The Sunday Age' August 1995). During November, the TAC commenced the development of new drink-driving advertisements to support the random breath test "booze bus" program which had been launched in September 1989 as part of the strategy. While the new advertisements were being developed for launching in mid-December, the TAC and VicRoads collaborated to place two existing VicRoads drink-driving television advertisements and three radio advertisements in the mass media over a 3-4 week period during November 1989. The television advertisements were placed at relatively high levels compared with previous practice, achieving 430 TARPs for audiences in the 16-24 year old age group and 530 TARPs for 25-54 year olds. The advertisements were recalled by 80-85% of these groups.

MUARC considered that it would have been inappropriate to ignore the drink-driving advertising during November 1989, given its relatively high intensity and recall, in the models for HAH crashes. The models were seeking to find any relationships between *levels* of drink-driving advertising and crashes, and were not concerned about the specific *content*. In RN52 (p. 18) the MUARC researchers declared that the VicRoads-developed advertisements had been used in November 1989. It would have been a greater error to totally ignore the drink-drive advertising during November 1989 because it had the potential to add awareness of drink-driving issues to that achieved by the TAC-developed advertisements launched in December 1989.

Systematic information on the levels of road safety advertising placed by VicRoads and the Federal Office of Road Safety prior to November 1989 was not readily available for consideration in conjunction with the television TARPs data which has been collected regularly since that time. There was no arbitrariness in the decision not to include road safety advertising levels prior to November 1989; the information was not available; and the November 1989 TARPs were included for the reason given above.

White et al (2000b) have also re-analysed the HAH serious casualty crash models in RN74 in which the November 1989 Adstock was ignored (set at zero) and the drink-driving advertising assumed to commence in December 1989. They found that the advertising variable was no longer statistically significant and then stated "If the December starting date had been used by Newstead et al (1995) [RN74], they would have had to conclude that drink drive advertising had no effect on HAH crash numbers". This statement reflects an apparent misunderstanding by White et al (2000a,b) about the aim and role of RN74; it was not to evaluate any one of the Victorian road safety initiatives, it was to rationalise and combine the results of prior evaluations to assess the overall contribution of each initiative. In addition, the statement is hypothetical because, as stated above, MUARC had information that relatively high levels of drink-driving advertising with high levels of recall had occurred during November 1989.

Notwithstanding this response, MUARC has repeated White et al's (2000b) analysis on the HAH serious casualty crash models developed in RN52 for which, as pointed out above, there were important differences in the treatment of the explanatory factors. When the November 1989 advertising was ignored, the elasticity of drink-driving Adstock dropped from –0.0249 ($p < 0.001$) to –0.0169 ($p = 0.008$), but it was still statistically significant. Thus MUARC's analysis appears to be in substantial conflict with White et al's (2000b) suggestion that the TAC drink-driving advertising from December 1989 was ineffective. Furthermore, MUARC's analysis is relevant to the findings of report RN52 (unlike White et al's analysis) which in turn were the basis of the economic analysis requested by TAC regarding road safety advertising levels.

In summary, MUARC had non-arbitrary reasons for the inclusion of levels of drink-driving advertising during November 1989 in the data used to develop the statistical models in RN52 and RN74. MUARC's inability to include advertising levels prior to November 1989 was due to the absence of relevant information in useable form. MUARC disagrees with White et al (2000b) about the criticality of the November 1989 data to MUARC's modelling results. MUARC did not find this to be the case when it re-analysed the model developed in RN52, which is the only MUARC report relevant to the question of road safety advertising levels.

WHITE ET AL'S CONCLUSIONS AND RECOMMENDATIONS

White et al (2000b) state "From the re-analyses of the data of Newstead et al (1995) [RN74] it is concluded that the … estimates of crash and financial savings attributable to the TAC-funded countermeasures are not supportable".

In response, MUARC re-iterates that RN74 is not representative of the MUARC research on enforcement and advertising in Victoria during the 1990's. Scientific evaluation studies, especially RN38 and RN42, have shown substantial reductions in road trauma associated with the RBT "booze buses" and the new speed cameras, each supported by TAC advertising.

RN74 was not considered to be a scientific evaluation of these programs. It aimed to combine the results of prior evaluations to assess the relative contributions of each program. Best evidence that the road trauma reductions are attributable to these programs is given in MUARC's other evaluation research.

White et al (2000b) also state "More particularly, the re-analyses [of RN74 data] have failed to support the claims of Cameron et al (1993) [RN52] concerning the crash reductions that can be achieved through high levels of TAC-funded road safety TV advertising".

In response, MUARC re-iterates that the data and analysis in RN74 is not representative of that in RN52. There were important differences in the objectives, data analysed, and assumptions of these two studies. RN52 provided estimates of the point of diminishing returns for TAC advertising, subject to stated assumptions of the analysis, whereas RN74 did not. The assumptions of RN52 were tested to a limited extent in the original research, and have since been tested by White et al and found to be satisfactory. White et al's findings from the re-analysis of the data in RN74, as well as being questionable in themselves, are not relevant to RN52.

White et al (2000b) recommend that "The developers of an innovative road crash countermeasure should always first consider the possibility of implementing the countermeasure in such a way that it can be evaluated experimentally".

MUARC supports this recommendation very strongly, however it notes that countermeasure developers have seldom implemented new initiatives in this way in the past. In practice, the implementers have been reluctant to withhold their initiatives from a part of the road user, vehicle or road environment populations so that an experimental design can be used for the evaluation. Ideally, the initiative should be withheld from a randomly-selected part of the eligible population, or from a part which has been matched with the treated sub-population on influential characteristics. This has seldom happened in Victoria and MUARC has been forced to use quasi-experimental designs, in which control crashes have been defined *post hoc* for comparison with treated crashes, in its evaluations of the enforcement and advertising programs. However, MUARC recognises that evaluation on this basis is not ideal, but represents the best available approach in non-experimental settings.

White et al (2000b) also recommend that "MUARC consultancy advice in favour of very high levels of road safety advertising should not be taken into consideration when determining appropriate levels of such advertising".

White et al make this recommendation on the basis of their re-analysis of the data in RN74 with "The main aim … [being] to check the soundness of the advice [in RN102] concerning the effectiveness of high levels of television advertising" (White et al 2000a,b). MUARC re-iterates a number of points made earlier with respect to its consultancy advice and the link between TAC road safety advertising and crashes:

- RN38 and RN42 provided scientific evidence of reductions in severe crashes associated with:
  - the RBT "booze buses", supported by TAC drink-driving advertising
  - the new speed cameras, supported by TAC speed-related advertising

- MUARC's macro-level trend analysis showed that the decrease in the Victorian road toll during 1990 and subsequent years was a greater reduction than that expected from pre-existing trends (this provides external support for the findings in RN38 and RN42)

- RN52 estimated the points of diminishing returns, under stated assumptions, for levels of drink-driving and speed-related TAC television advertising, respectively

- the assumptions made by MUARC in RN52 have been tested and found to be satisfactory

- the advice provided to the South Australian Office of Road Safety in RN102 made reference to RN38, RN42 and RN52 (but not RN74) and to data which showed that the level of road safety advertising in Adelaide was less than half the level in Melbourne

- the advice was not dependent on the findings in RN52 (had the advice been based on the points of diminishing returns, the recommendation would have been to more than double the level of advertising in South Australia)

- the statistical models of monthly crash variations developed in RN52 were not based on "data dredging" to find enforcement, advertising and socio-economic variables, which then may have had only spurious relationships with the crashes, to include in the models

- when White et al's (2000b) CMA 12 Leading Index replaced the unemployment rate in the RN74's LAH serious casualty crash model, the speed-related enforcement and advertising variables remained as statistically significant factors explaining the crash variations

- speed-related advertising was a statistically significant factor in White et al's (2000b) CMA 12 Leading Index model of monthly serious casualty crashes during all times of the week (this finding also indicated that White et al's three-factor model is not parsimonious)

- White et al's (2000b) test of whether the enforcement and advertising variables in the RN74 models were quantitatively related to the crash variations is not valid and does not establish the absence of quantitative relationships

- White et al's (1999) earlier investigation of the micro-level effects on LAH serious casualty crashes found that the speed-related advertising was statistically significant (and the speed-related enforcement nearly significant)

These points indicate that the basis for MUARC's consultancy advice is sound. Based on White et al's investigations, the evidence in favour of the effectiveness of TAC speed-related advertising supporting the Victorian speed camera program has been strengthened. The evidence for the effects of the drink-driving advertising supporting Victoria's RBT program relies on RN38 and on relevant parts of RN52 (which White et al have tested and found satisfactory).

MUARC'S EARLIER RESPONSE TO WHITE ET AL'S INVESTIGATIONS

A general response to the January 1999 draft report was prepared by MUARC in March 1999. A discussion was held with representatives of the South Australian Office of Road Safety during April 1999.

A number of issues raised in MUARC's general response have not been adequately addressed in the November 2000 report (White et al 2000b). These include:

- the re-investigation should be focused on the data and analysis used in RN52, not RN74, because RN52 is relevant to the aim of the re-investigation stated by White et al, whereas RN74 is not
- MUARC's objection to White et al's claim that MUARC portrayed relationships between advertising levels and crashes as "dose-response" relationships
- White et al's failure to include MUARC's multivariate explanation for the timing of the change in crash trends, and for the different peaks related to the severity of the crash series considered.
- White et al's failure to include the findings that diagnostic tests of the models in RN52 were satisfactory
- White et al's failure to include their own findings of evidence of micro-level effects of the TAC speed-related advertising
- MUARC's objection to the use of "data dredging" to find a replacement for unemployment rate as an indicator of economic conditions
- White et al's failure to include MUARC's full explanation for the inclusion of the November 1989 advertising with the later data
  (MUARC's re-analysis of the RN52 data, which indicated that this was not a critical decision, has been described in a token manner by White et al 2000b)

MUARC'S CONCLUSIONS

1. Scientific evaluations conducted by MUARC have shown substantial reductions in road trauma in Victoria due to increased random breath testing using "booze buses" and the new speed camera program, each supported by TAC advertising.

2. The statistical models of monthly casualty crashes as functions of enforcement, advertising and socio-economic factors, developed in RN52, are sound. They have been tested by MUARC and by White et al and have been found to be satisfactory. White et al's investigations have provided additional evidence of the relationship between the TAC speed-related advertising and crashes.

3. The estimates of the points of diminishing returns of levels of TAC drink-driving and speed-related advertising, originally provided in RN52 based on the statistical models in that report, are sound. The economic analysis of advertising levels, which was based on the coefficients of the advertising variables in the statistical models, has not been questioned.

4. White et al's re-analysis of the data used in RN74 is not relevant to RN52 because of important differences in their objectives, the types of crashes analysed, the time periods covered, the treatment of levels of speed-related advertising, the inclusion of car-based random breath testing, and the assumptions made and subsequently tested.

5. White et al's so-called parsimonious three-factor model of crash variations in Victoria was based on data dredging and cannot be considered to be a valid alternative to MUARC's models.

6. White et al's test of the quantitative relationships between crashes and the enforcement and advertising variables is not valid. It is not an adequate test of the presence or absence of quantitative relationships.

REFERENCES

BOX, GEP, and TIAO, GC (1975), "Intervention analysis with applications to economic and environmental problems". *Journal of the American Statistical Association*, Vol. 70, *349*, pp. 70-79.

BROADBENT, S (1979), "One way TV advertisements work". Journal of the Market Research Society, Vol. 21, No. 3. London.

CAMERON, MH (1997), "Comparison of methods of modelling trends in road casualties in Australasia and The Netherlands". Proceedings, Road Safety Research and Enforcement Conference, Hobart, Tasmania.

CAMERON, MH, and VULCAN, AP (1998), "Evaluation review of the Supplementary Road Safety Package and its outcomes during the first two years". Report to Land Transport Safety Authority, New Zealand.

CAMERON, MH, HAWORTH, N, OXLEY, J, NEWSTEAD, SV, and LE, T (1993), "Evaluation of Transport Accident Commission road safety television advertising". Report No. 52, Monash University Accident Research Centre.

CAMERON, MH, NEWSTEAD, S, and VULCAN,AP (1994) "Analysis of reductions in Victorian road casualties, 1989 to 1992". *Proceedings 17th ARRB Conference,* Part 5, pp 165-182.

COUNCIL, FM, REINFURT, DW, CAMPBELL, BJ, ROEDIGER, FL, CARROLL, L, DUTT, A, and DUNHAM, JR (1980), *Accident Research Manual.* Highway Safety Research Center, University of North Carolina, NC.

ECCLESTON, J (2000), "A Peer Review of the Transport SA Report 'Re-investigation of data on the effectiveness of the Victorian Transport Accident Commission's road safety television advertising campaigns' ". Department of Mathematics, University of Queensland, Brisbane.

FRY, TRL (1996), "Advertising wearout in the Transport Accident Commission road safety campaigns". *Accident Analysis and Prevention,* Vol. 28, No. 1.

GARBER, S, and GRAHAM, JD (1990), "The effects of the new 65 mile-per-hour speed limit on rural highway fatalities: A state-by-state analysis". *Accident Analysis and Prevention*, Vol. 22, No. 2.

GRAHAM, A, and TAYLOR, R (1994), "Road Safety 2000: Trends and targets". *Proceedings, Road Safety 2000 Review Conference*, Sydney. Roads and Traffic Authority, New South Wales, Australia.

HAKIM, S, SHEFER, D, HAKKERT, AS, and HOCHERMAN, I (1991), "A critical review of macro models for road accidents". *Accident Analysis and Prevention*, Vol. 23, No. 5.

HARRISON, WA (1990), "Update of alcohol times as a surrogate measure of alcohol-involvement in accidents". Research Note, Monash University Accident Research Centre.

HARRY, A (1997), "Economic indicators in road crash modelling". Safety Strategy, Transport South Australia.

LAMBERT, J (1992), "Fuel consumption in Victoria and estimates of vehicle travel". Report No. GR92-5, VicRoads, Victoria.

LAND TRANSPORT SAFETY AUTHORITY (1996), "A road safety resource allocation model". Working paper 1, LTSA, New Zealand.

LAND TRANSPORT SAFETY AUTHORITY (1998a), "The *Safety Directions* Development Programme". Working paper 2, LTSA, New Zealand.

LAND TRANSPORT SAFETY AUTHORITY (1998b), "An international comparison of road safety enforcement". Working paper 3, LTSA, New Zealand.

LAND TRANSPORT SAFETY AUTHORITY (1998c), "Setting road safety targets". Working paper 4, LTSA, New Zealand.

NEWSTEAD, SV, CAMERON, MH, GANTZER, S, and VULCAN, P (1995), "Modelling of some major factors influencing road trauma trends in Victoria 1989-93". Report No. 74, Monash University Accident Research Centre.

OPPE, S (1989), Macroscopic models for traffic and traffic safety. *Accident Analysis and Prevention*, Vol.21, No. 3, pp. 225-232.

OPPE, S (1991a), The development of traffic and traffic safety in six developed countries. *Accident Analysis and Prevention*, Vol. 23, No. 5, pp. 401-412.

OPPE, S (1991b), Development of traffic and traffic safety: global trends and incidental fluctuations. *Accident Analysis and Prevention*, Vol. 23, No. 5, pp. 413-422.

SHTIFELMAN, M, DIAMANTOPOULOU, K, and CAMERON, MH (1999), "Measuring drink-driving advertising awareness in Victoria". Proceedings, Road Safety Research, Policing and Education Conference, Canberra.

SMEED, RJ (1968), Some statistical aspects of road safety research. *Journal of the Royal Statistical Society, Series A (General)*, Vol. 112, No. 1, pp.1-34.

'THE SUNDAY AGE' (1995), "A crash course in road safety". View magazine, 20 August – 26 August 1995.

VULCAN, AP (1993), "The road toll in Victoria - An objective analysis", Proceedings, Road Safety Forum, Melbourne.

VULCAN, AP, and CAMERON, MH (1996), "Review of road safety advertising in New Zealand". Report to Land Transport Safety Authority, New Zealand.

VULCAN, AP, CAMERON, MH, MULLAN, N, and DYTE, D (1996), "Possibility of adapting some road safety measures successfully applied in Victoria to South Australia". Report No. 102, Monash University Accident Research Centre.

WAGENAAR, AC, STREFF, FM, and SCHULTZ, R.H. (1990), "Effects of the 65 mph speed limit on injury morbidity and mortality". *Accident Analysis and Prevention*, Vol. 22, No. 6, pp. 571-585.

WHITE, M, WALKER, J, GLONEK, G, and BURNS, N (2000a), "Re-investigation of the Effectiveness of the Victorian TAC's Road Safety Campaigns". Paper submitted to Road Safety Research, Policing and Education Conference, Brisbane, November 2000.

WHITE, M, WALKER, J, GLONEK, G, and BURNS, N (2000b), "Re-investigation of the Effectiveness of the Victorian Transport Accident Commission's Road Safety Campaigns". Report No. 4/2000, Safety Strategy, Transport South Australia, November 2000.