C. Boughton

DEPARTMENT OF TRANSPORT

THE USE OF TELEVISION PUBLICITY TO MODIFY SEAT BELT WEARING BEHAVIOUR

I. R. Johnston and M. H. Cameron

Australian Government Publishing Service
Canberra 1979

ISBN 0 642 01567 8

Document Retrieval Information

Date: April 1979

Keywords

Safety belt adjustment; safety belt usage; safety education; campaigns/public information programs/publicity; television (media)

Abstract

Legislation requiring seat belts to be worn has led to major reductions in casualties in Australia. However injuries have resulted from belt slackness, from the location of buckles in the abdominal region and from webbing twist. The design changes which have been effected to minimise the occurrence of these problems will take many years to permeate the vehicle population and as an interim measure a television publicity campaign was conducted to encourage occupants to better adjust their belts.

Three experiments to evaluate the effectiveness of the campaign are described. Using a criterion measure of change in belt adujstment observed during roadside surveys, the experiments demonstrated that 'intense' exposure to television publicity over a 'short' period achieved significant decreases in the incidences of loosely adjusted belts, the location of buckles on the abdomen, and twist in belt webbing. The exposure involved three to four screenings per night during peak time on each available commercial station for fourteen consecutive nights. Other combinations of intensity and duration which were tested did not yield consistent positive results.

The study establishes the value of television publicity as a countermeasure under certain specific conditions. It also demonstrates that valid research to measure 'real world' publicity effects is feasible.

Note

This report is disseminated in the interest of information exchange.

The views expressed are those of the author(s) and do not necessarily represent those of the Commonwealth Government.

Commonwealth Department of Transport
Office of Road Safety
Box 1839Q G.P.O.
Melbourne, Vic. 3001

M. H. Cameron & Associates
17 Myrtle Grove
Blackburn, Vic. 3130

Printed by C. J. THOMPSON, Commonwealth Government Printer, Canberra

Contents

24702/79—2

Introduction

The problem

Between 1970 and 1972 legislation requiring vehicle occupants to wear seat belts came into force in all Australian States and Territories. Although it is arguably the most successful single countermeasure introduced to date in Australia its full potential has not been realised.[9] The major degrading factors are that not all vehicle occupants have seat belts available to them; of those that do, the wearing rate is less than 100 per cent; and of those that wear their belts not all have them optimally adjusted. It is to the last of these factors that the research described in this report was directed.

As part of an extensive research program to assess the effectiveness of the legislation the Department of Transport commissioned an in-depth study of injuries sustained in crashes where seat belts had been worn. By using a specially designed measuring jig, Ryan and Baldwin were able retrospectively to estimate seat belt adjustment in the crash vehicles.[14] They concluded that, in many cases, injury was associated with serious maladjustment. The major maladjustments were excessive belt slack and the buckle being worn on or forward of the hip. These findings were later confirmed by other studies.[6, 7]

Ryan and Baldwin then conducted a small-scale roadside survey and estimated that only about half of the seat belts observed were satisfactorily adjusted. In an earlier roadside study (late 1971) Andreassend classified only 14 per cent of drivers as having their belts satisfactorily adjusted.[1] The variation between the estimates may have been due to differences in the operational definitions of satisfactory adjustment, small sample sizes, the timing of the studies or to other methodological differences.

Nevertheless, as early as the end of 1972 it was clear that large proportions of seat belt wearers, possibly a majority, were wearing their belts in less than an optimum manner and that avoidable injuries were occurring. It should be stressed that Ryan and Baldwin concluded from their study that the injuries received would have been more severe had belts not been worn.[14]

In the absence of detailed and reliable data on both injuries and seat belt wearing in the police-reported mass data system, the magnitude of the problem could not be measured. However, given the indications of widespread maladjustment in the at-risk population, there seemed little doubt that it was a serious problem.

Possible countermeasures

The advent of laws requiring seat belts to be worn brought about a dramatic increase in wearing rate; within a year of the law becoming effective in Victoria the wearing rate of available belts for drivers and front left passengers in metropolitan Melbourne had risen to over 75 per cent.[1]

For several reasons a high incidence of maladjustment among this new population of belt wearers was not surprising. Many—almost one quarter of those interviewed by Andreassend—were fastening their belts only because it was legally required.[1] Although, technically, the law required a belt to be 'properly adjusted and securely fastened', education programs accompanying the new law rarely mentioned manner of adjustment and it seems reasonable to assume that large numbers of vehicle occupants neither appreciated the need for proper adjustment nor understood how to achieve it. Further, most of the three-point belt systems in common use required considerable effort to adjust optimally; webbing being subject to twist and both buckle webbing and torso webbing typically requiring separate adjustment for length.

Two countermeasures appeared to be required. Improvements to the design and installation of seat belt systems were necessary to facilitate—and ideally to remove the need for—individual manual adjustment and education was needed to inform seat belt users of the need for proper adjustment and of the methods of achieving it.

Progressive improvements to seat belt systems have been achieved through amendments to the relevant vehicle safety design rules, the first becoming effective for passenger vehicles manufactured after 1 January 1974.[9] Inertia reel belts—which remove the need to adjust for tightness and which also minimise webbing twist—are now mandatory fittings in the front outboard seating positions of new passenger cars. Moreover, current seat belt systems are designed for one-handed operation and, through the use of a buckle stalk or through limitations on the length of buckle webbing, have largely overcome the risk of locating the buckle forward of the hip.

Nevertheless, the progressive introduction in new vehicles of these design changes has meant that a large and relatively slowly decreasing proportion of the vehicle population has seat belts requiring considerable manual adjustment. An education program to inform and assist the users of these vehicles was considered to be a necessary holding measure.

Potential effectiveness

Several features of the seat belt maladjustment problem combined to create both a reasonable expectation of success for publicity as a countermeasure and conditions favourable to a controlled evaluation of effectiveness:

- the basic behaviour—seat belt wearing—was widespread, publicly accepted and its safety value widely acknowledged[1, 4];
- the incidence of the target behaviour—optimal seat belt adjustment—was relatively low, giving ample scope for change[1, 14];
- there were reasonable grounds for assuming that one important reason for this low incidence was public ignorance;
- the target behaviour was highly specific and capable of being demonstrated unambiguously;
- target behaviour could be performed at relatively little cost by the target group, within the limits of inconvenient seat belt designs;
- the target behaviour could be objectively measured;
- the target behaviour was known to be directly related to injury causation.

The principles underlying these features are those considered essential to the success of a publicity program.[10, 11] For this reason it was important that research be undertaken to evaluate effectiveness.

Development of the publicity program

Target behaviours

The ultimate aim of the publicity program was to reduce the incidence of injuries related to poor seat belt adjustment; the sub-goal, or mechanism, was to increase the proportion of belt wearers with optimally adjusted seat belts. Three specific target behaviours were chosen—degree of tightness, location of buckle and webbing twist. The first two were selected because of the established direct role of slackness and location of the buckle on or forward of the hip in injury causation[14] and the third because of its potential role; significant webbing twist prevents the distribution of crash loads in the manner for which the restraint was designed.

Target group

Neither of the two previous surveys of manner of adjustment provided details of the incidence of the target behaviours as functions of occupant seating position, age, sex or other relevant parameters.[1, 14] In view of the very low proportions of drivers rated as having their belts satisfactorily adjusted it was decided that the target group should comprise all occupants of passenger cars and derivatives to whom the compulsory wearing legislation applied.

Choice of medium

With a broadly defined target group and with target behaviours which required an active visual demonstration if they were to be conveyed clearly and unambiguously, television was the obvious primary medium. Press advertising, and other printed material (pamphlets, posters etc.), although visual, lacked the element of action considered essential to communicating the target behaviours.

Nature of the communication

A sixty second black and white television film was produced. In sequence, it showed an ambulance pulling into a hospital casualty department and a 'victim' being wheeled in; a surgeon 'washing up', who reinforces the value of seat belts but reflects on the occurrence of needless injuries through poorly adjusted belts; the surgeon entering his car in the hospital car park; and a detailed demonstration by the surgeon of optimal adjustment—locating the buckle rearward of the hip, ensuring there is no twist in the webbing, fastening the buckle and then removing all slack.

A four page pamphlet and a poster were produced as supplementary publicity material. Both used the surgeon from the television film and illustrated the three target behaviours.

Several aspects of the approach used require comment. In keeping with recommended practice the approach was a highly specific, concrete demonstration of the target behaviours performed by an authority figure having high credibility and status.[2, 11, 16] Of the communication principles expounded by Wilde[16]—specificity of

24702/79—3

message, concrete instructiveness, personal relevance, ability to facilitate imitative behaviour and immediacy—only the immediacy criterion was not met; at the time of receiving the message the recommended behaviour could not be performed.

The use of a mild fear setting as the motivating factor also requires justification. There is a reasonable body of evidence which indicates that fear can be a successful motivator if the behaviour to be changed is not well established and can be changed without great cost by the target group, and if the recommended behaviour is specified in detail and can be implemented rapidly.[2, 11] The seat belt adjustment target behaviours satisfy each of these conditions.

The main alternative motivating factor was to stress the positive benefits of a well-adjusted belt; feelings of security, comfort and so on. However, given the known difficulties of easily achieving both comfortable and optimal adjustment it was considered that such an approach would have been likely to create a credibility gap between communication and target group.

Development of the research design

The ideal experiment

In the ideal experiment, car occupants would be randomly assigned to treatment and control groups; random samples from each group then being chosen and each individual being measured before and after the treatment group is exposed to the publicity.[5] (Figure 1)

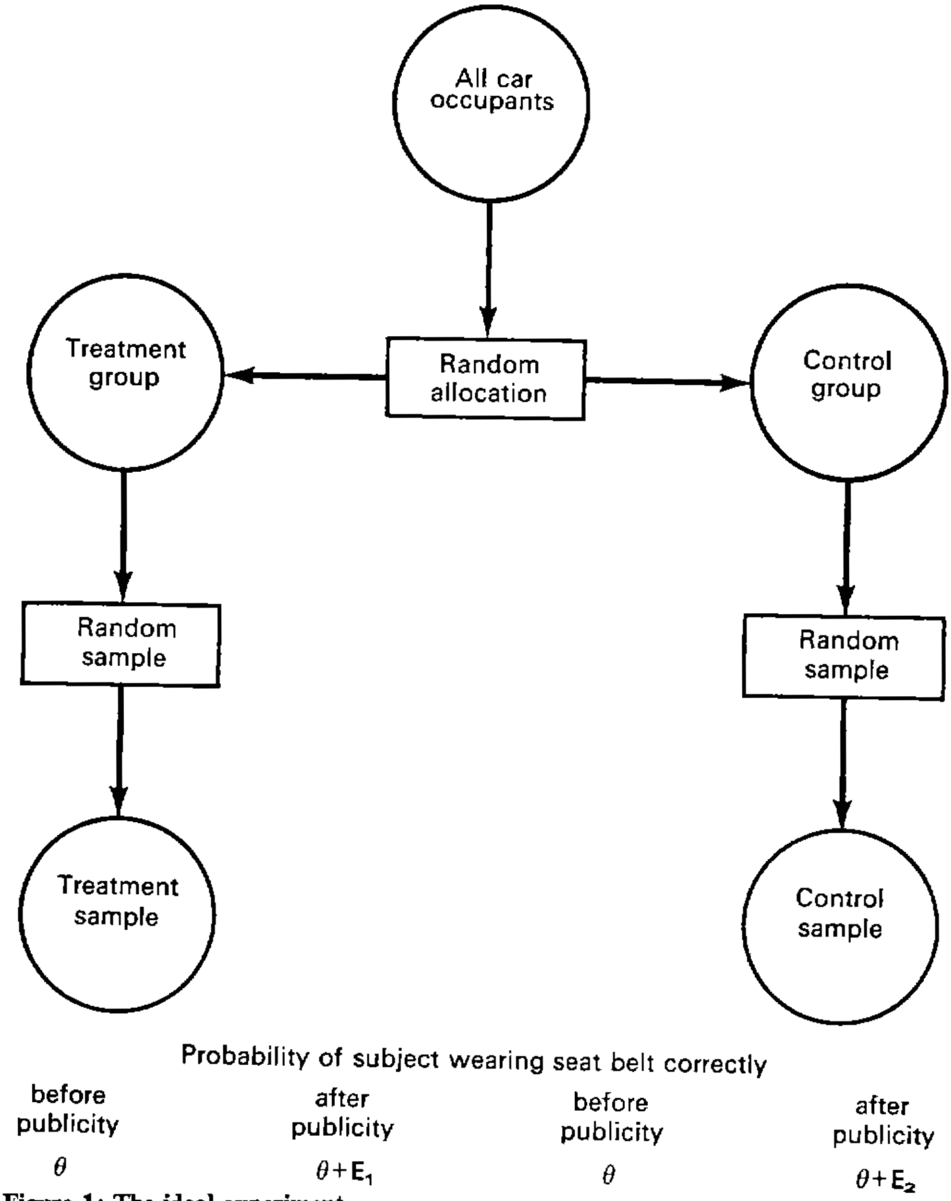


Figure 1: The ideal experiment

The effect of the publicity is measured by $E_1 - E_2$. E_2 is not necessarily equal to zero as the probability of people in the control group wearing their seat belts correctly could change due to other stimuli or trend effects. The treatment group is similarly exposed to these base level effects, plus the publicity.

The ideal experiment has the following key requirements:

- individuals are randomly assigned to the treatment or control group (or the decision that each particular person receives the treatment is a random one) and
- individuals are identifiable as to whether they belong to the treatment or control group (so that this can be noted at the times of measurement).

These requirements are difficult to achieve when the treatment is publicity using one of the mass media. With television the decision whether each person is exposed to the medium or not is not a random one. Moreover, it is very difficult to identify those people who received the publicity treatment except by direct questioning, which tends to violate the need for non-reactive measurement.[5]

An experiment conducted in the United States of America by Robertson *et al* comes close to satisfying the requirements of the ideal experiment.[13] They had access to a dual cable television system designed for marketing studies. The two cables were distributed in a checkerboard fashion among blocks of households that had chosen to pay for the improved signal which the cable provided. Although the assignment of households to one or another cable was not strictly random, various marketing studies found no significant differences between the two groups of households as far as many socio-economic characteristics were concerned.

The treatment was television publicity designed to encourage viewers to wear seat belts when driving. Observations of car occupants were made unobtrusively to determine their seat belt wearing before and during the publicity campaign. The registration number of the car was also noted, which allowed identification of the household owning it and hence whether the occupants belonged to the treatment group or one of the two control groups; households on the other cable and households on neither cable.

Experimental design adopted

The absence of cable television in Australia prevented the use of an experimental design like that used by Robertson *et al.*[13] The treatment group could only be taken as all the people living within the area of coverage of a city's television stations, whether or not they viewed the publicity. The control group was taken as the people living in another city, selected so that the two cities were sufficiently far apart for their areas of television coverage not to overlap. Note that this design requires the assumption that car occupants travelling in a city view only the television of that city.

The assignment of treatment and control status to the two cities, whether at random or not, cannot be considered a random assignment of the treatment to the pooled group of people living in the two cities. These two groups of people will differ by the characteristics which cause them to live in different cities, as well as by the effects of the different environments in which they live. This difficulty was overcome by conducting two experiments in which the roles of the treatment and control cities were reversed in the second experiment.

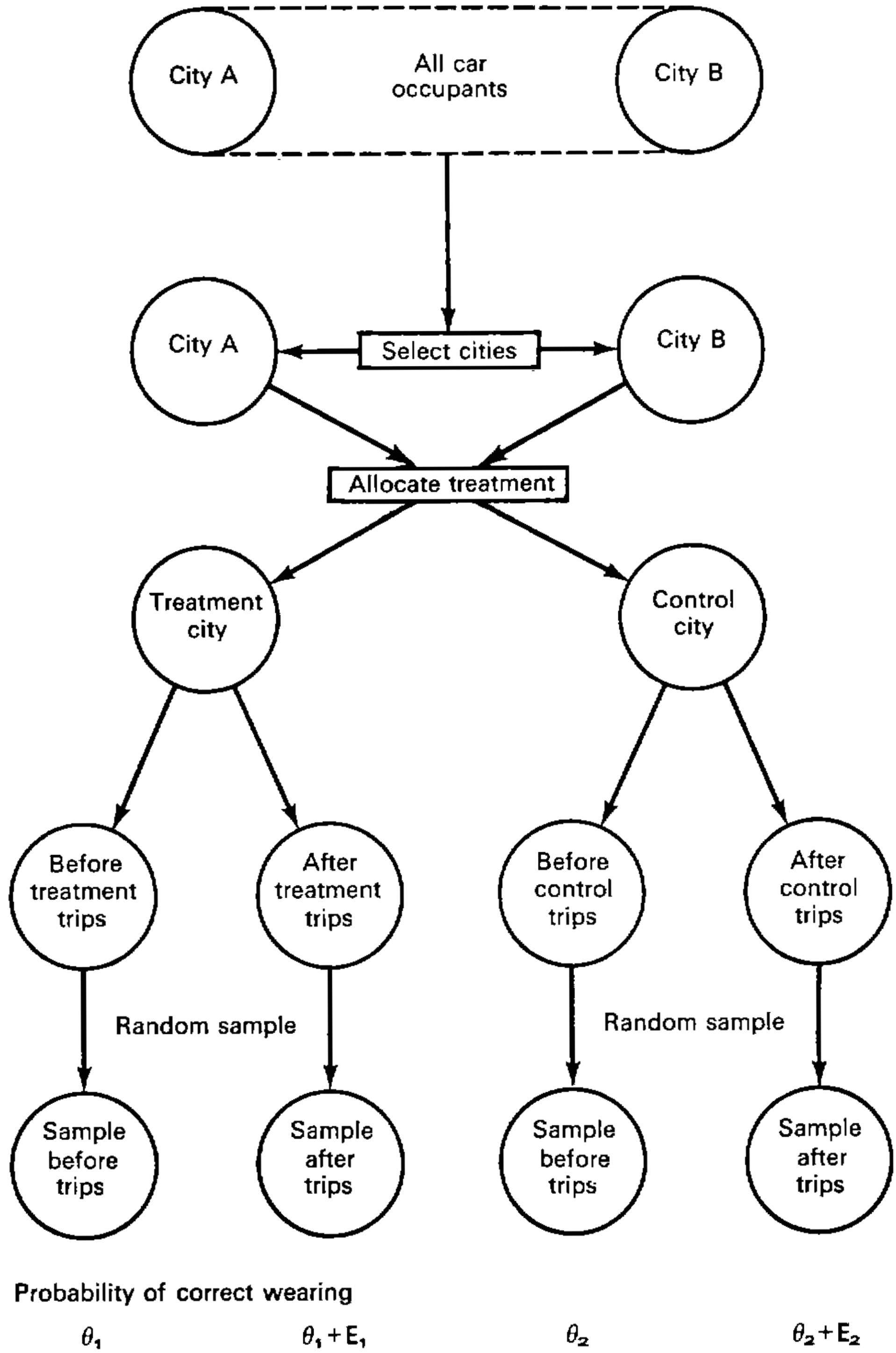


Figure 2: Experimental design used

The design used is illustrated in Figure 2 for comparison with Figure 1. Note that not necessarily the same people make trips after the publicity as before it, nor need the random samples from these two groups include the same people. This contrasts with the ideal experiment, where the samples become fixed groups to be observed both before and after. This departure from the ideal experiment also occurred in the study conducted by Robertson *et al.*[13] It is only a problem if the treatment is likely to cause subjects to change their travel behaviour. Although the publicity under evaluation has a theme related to road travel, such a change is unlikely.

Note also that the probability of an occupant wearing his seat belt correctly before the publicity is not necessarily the same in the treatment (θ_1) and control (θ_2) cities. These probabilities are the same (that is, θ) in the ideal experiment because of the random allocation of people to the treatment and control groups. The variation between θ_1 and θ_2 affects the interpretation of E_1 and E_2, and the effect of the publicity could be measured by $E_1 - E_2$, the relative change, or by $E_1/\theta_1 - E_2/\theta_2$, the relative proportionate change.

After the cities were selected and the treatment publicity allocated, efforts were made to avoid competing publicity in the control city. Members of the Publicity Advisory Committee on Education in Road Safety* were asked to defer similar or related publicity in the study cities. However, no similar control was possible over concurrent events such as increased enforcement of seat belt wearing in the treatment or control city, or publicity sponsored by other sources. The fact remains that treatment and control status were arbitrarily assigned to people living in different environments, exposed to different changes in their environment concurrent with the publicity campaign in the treatment city.

* A federal-state committee which co-ordinates national road safety publicity programs under the auspices of the ministerial Australian Transport Advisory Council. The assistance of committee members is gratefully acknowledged.

Method

Measurement criteria

Since the ultimate goal of the program was to reduce the incidence of injuries associated with poor seat belt adjustment the most valid criterion measure is injuries to seat-belted vehicle occupants. Unfortunately seat belt wearing by persons in crashes is not reliably reported and moreover the extent of detailed injury data that would be needed is not routinely available.[14] However, in view of the demonstrated direct causal link between the target behaviours and injury occurrence, behaviour change is not only itself a valid criterion but has the added advantage of being a more sensitive measure.

The publicity demonstrated optimal adjustment only for a three-point static belt. This type of belt was chosen because it was almost universally fitted to the front outboard seating positions of 1969 and later model vehicles (at least up to the time of the study). Following pilot observations operational measures were developed for each of the target behaviours. They were:

- *Tightness*
  - the occupant was asked to lean forward as far as he could;
  - 'tight'—where occupant had virtually no forward movement of body (during training this was related to the passage of a clenched fist between sash and chest, during observation a physical measure could not, of course, be employed);
  - 'in-between'—where movement was slight to moderate (equivalent to the passage of a laterally extended hand between sash and chest);
  - 'loose'—when movement was greater than either of the above.

- *Buckle position*
  - 'off-body'—where buckle was by occupant's side and rearward of the hip joint;
  - 'in-between'—where buckle was in the forward region of the hip area;
  - 'too near middle'—where buckle was on abdomen.

- *Twist*
  - 'twist'—where sash webbing was twisted whilst in contact with body (twist in lap webbing was not considered because it was too difficult to observe reliably);
  - 'no twist'.

Teams of observers were trained thoroughly. Training consisted of approximately half a day in the classroom undergoing a detailed briefing on the total measurement procedure, including a verbal and pictorial demonstration of the criteria; a session of about an hour making assessments of controlled belt settings using a stationary car; and a further half-day session in the field making practice assessments under close supervision.

Method of measurement

It was extremely difficult to develop a relatively unobstrusive, non-reactive method of measuring behaviour representing a random sample of vehicle occupant trips in which seat belts are worn and several practical compromises were necessary. The major features of the technique were:

- Eligible vehicles comprised all passenger cars and derivatives;

  observations were restricted to the driver and front outboard passenger (if any)—a maximum of two occupants per vehicle.

- To ensure that measurement was as naturalistic as possible and to avoid disrupting traffic flow, all observations were made whilst vehicles were stationary during the red light phase at signal-controlled intersections. These phases varied between twenty-five and forty seconds in duration.

- The technique of utilising natural traffic pauses necessitated sampling traffic only at signal controlled intersections. To ensure his safety the observer was stationed on the central median strip and was not permitted to walk on the carriageway. This requirement limited potential sites to signal controlled intersections on arterial roads divided by a central median. Individual sites were arbitrarily chosen to maximise observer safety, to maximise the number of observations per unit time, and to cover a range of locations within each city with largely independent traffic streams.

- The sampling technique was simple; the observer recorded data for the occupants of the first eligible vehicle stopping at the red light. This provided a reasonable approximation to a random sample of vehicles travelling on arterial roads.

- Observations were made between the hours of 6 a.m. and midnight (Experiments 1 and 2; 11 p.m. in Experiment 3) on the four days (Thursday to Sunday) immediately preceding the commencement of publicity and on the first occurring corresponding four days at the conclusion of the publicity. Five to six sites, chosen arbitrarily on the above grounds, were used in each city. Measurement was not continuous at each site. A wide range of both times of day and days of week was sampled to ensure measurement of a reasonable representation of the target behaviours.

Data collected

The following data were collected for the driver and front outboard passenger (if any) of each vehicle observed.

- Whether a three-point static belt was fitted (lap only, sash only and full harness were recorded as three-point not fitted);

- If fitted, whether the three-point belt was
  - — buckled,
  - — draped over shoulder, but not buckled,
  - — left hanging from pillar,
  - — inertia reel;

- If buckled, assessments were made of
  - — tightness,
  - — buckle position,
  - — twist;

  using the operational definition described earlier.
- Sex and estimate age (8–29, 30–49, 50 +) of the occupant.
- Time of day, day of week, site. (Prevailing weather conditions and wearing/not wearing overcoat were additional descriptors in Experiment 3.)

Method of analysis

The statistical technique employed to analyse the data was the three-way chi-square test for second-order interactions in complex contingency tables for three categorical variables.[6] The three categorical variables were city (treatment(s) and control), time (before and after publicity) and measures of correctness of adjustment.

The detailed development of this statistical technique, together with the rationale for calculations of the number of observations required to ensure sufficient statistical power, are given in Appendix A.

24702/79—4

Experiment 1

Method

The first experiment took place in the period May to July 1973. Three cities, each a State capital, were used and roadside observations of the manner of seat belt adjustment were made immediately before and after the television film was screened.

Hobart was the control city; screening of the film did not occur and distribution of the supplementary pamphlet and poster was withheld. Duration of the publicity program was held constant at fourteen days in the two treatment cities but intensity was varied. Adelaide received three screenings of the sixty-second film on each of its three commercial television channels for fourteen consecutive evenings, all in the prime advertising time period of 6–9 p.m. This 'high' intensity was selected on advice from advertising agencies as being roughly equivalent to that for the launching of a new product on the market. Melbourne received one screening on each of its three commercial channels per evening in the same time slot (Table 1).

Table 1: Experimental design

City	*Before measurements*	*Publicity treatment*	*After measurements*
Adelaide . .	6 a.m.–Midnight 17–20 May at a minimum of five sites	HIGH (3 × 60 seconds per night per commercial television channel for 14 consecutive nights)	6 a.m.–Midnight 7–10 June at identical sites
Melbourne .	As above . .	LOW (1 × 60 seconds otherwise as above)	As above
Hobart . .	As above . .	CONTROL (Nil) .	As above

As outlined earlier and described fully in Appendix A the basic method of analysis was the three-dimensional chi-square test. This technique was also used to investigate whether there was a need for 'controlled' analysis, that is, whether the effect of the publicity needed to be tested for each of various subsets of the data to avoid spurious conclusions. Each vehicle occupant observed for seat belt adjustment can also be described by the following five descriptors:

- age (8–29, 30–49, 50+);
- sex;
- seating position (driver or left front passenger);
- time of day (three hour groupings were used);
- day of week (Thursday to Sunday only).

It is necessary to control for occupant type when investigating the effect of the publicity on a particular measure of correctness if both:

- the apparent effect of the publicity in the treatment city varies significantly by occupant type (for example, women may improve tightness more than men) and
- the relative proportion of the occupants in this descriptive category varies significantly between the before and after stages in *either the treatment or control city* (for example, a higher proportion of women may be observed in the after stage).

In other words, a spurious conclusion could arise in a non-controlled analysis if there was a particular type of occupant who was heavily affected by the publicity and who was over-represented in the after stage in, say, the treatment city. In such a case there would appear to be a greater effect of the publicity than really existed for the average occupant.

The large number of controlled analyses led to a major problem of interpretation and presentation. A significant three-way interaction between each of the trichotomous measures of correctness (tightness and buckle position) and city and time was difficult to interpret. For example, an improvement in tightness in the treatment city relative to the control city could be due to:

- a drift from 'loose' to 'in-between'; or
- a drift from 'in-between' to 'tight'; or
- both.

A large number of significant changes of this type are difficult to summarise. This problem does not arise with the flatness measure as change can only be a drift from 'twist' to 'no-twist', or vice versa.

Accordingly, for ease of presentation, it was decided to re-define tightness and buckle position as dichotomies and carry out three-dimensional chi-square tests on these re-defined measures of correctness as well as the original trichotomous definitions.

Of the three levels of tightness, only 'loose' was considered a particularly dangerous condition. 'Tight' and 'in-between' were considered to be operationally rather stringent, so it was decided to pool these two levels to form the dichotomous tightness measure*.

For buckle position, a buckle in either the 'in-between' or 'near middle' position is capable of causing injury in a crash. Only 'off-hip' is a completely safe condition. Accordingly, it was decided to pool 'in-between' and 'near middle' to form the buckle position dichotomy.

Results

All occupants

Tables 2 and 3 summarise the results for all front seat occupants. (The results are given in full in Appendix B.) The HIGH *v.* CONTROL differences were highly significant for all three measures of correctness, indicating a real improvement in seat belt adjustment as a function of television publicity. The percentage wearing their seat belt 'tight' or 'in-between' increased by 8.7 per cent (from a base of 61.0

* These measures were now also consistent with Ryan and Baldwin's definition of tight/loose.[14]

per cent), 'no twist' by 6.3 per cent (from 75.0 per cent), and buckle 'off-hip' by 12.3 per cent (from 52.8 per cent). The LOW *v.* CONTROL differences were not significant for any of the target behaviours.

The HIGH *v.* LOW differences were highly significant for all three measures of correctness, confirming the real effect of the high level of publicity.

Table 2: Per cent change × target behaviour × publicity condition

	Publicity condition		
	HIGH (*Adelaide*)	*LOW* (*Melbourne*)	*CONTROL* (*Hobart*)
Per cent change in proportion 'tight' or 'in-between'	8.7	3.0	2.8
Number of observations before/after	3403/3420	3874/3107	2799/2097
Per cent change in proportion buckle 'off hip'	12.3	2.9	2.8
Number of observations before/after	3379/3407	3867/3105	2792/2088
Per cent change in proportion with 'no twist'	6.3	1.8	−0.1
Number of observations before/after	3381/3405	3865/3105	2793/2092

Table 3: Significance levels of statistical tests

	High v. Control	*Low v. Control*	*High v. Low*
Tightness	< 0.01	N.S.	< 0.001
Buckle position	< 0.001	N.S.	< 0.001
Twist	< 0.001	N.S.	< 0.001

Controlled analyses

Only the Adelaide data were investigated in detail to determine if there was need for controlled analysis. The apparent effect of the publicity on tightness was found to vary significantly with time of day (< 0.001) and age (< 0.05), and on buckle position with time of day (< 0.01). The relative distribution of observations by time of day and age also varied significantly between the before and after stages in Adelaide. No significant variation of the apparent publicity effect by day of week was found, but both Melbourne and Hobart had significant changes in the distribution of observations by day of week. Melbourne was not investigated in detail so it was not known if it was necessary to control for day of week. Accordingly, it was decided to adopt a conservative stance and control for time of day, age and day of week separately when analysing the effect of the publicity on each of the three measures of correctness.

The detailed results from the controlled analyses are given in Appendix B.

Decay effects

To measure whether the change in behaviour was maintained over time further measurements were taken, using the same procedure, six weeks after the publicity had ceased. Since there had been minimal change in Melbourne (LOW) compared with the control city of Hobart it was decided for economic reasons to use Melbourne as the control thereby avoiding the cost of further counts in Hobart.

Unfortunately, two events destroyed the validity of these follow-up observations. First, one of the three commercial channels in Melbourne had continued to screen the television film free of charge as a public service making Melbourne an impure control. Secondly, a high proportion of the original observers were unavailable and new observers had to be recruited and trained. Analyses of the results by individual observers revealed considerable differences that seemed to reflect different operational definitions of the criteria of correctness of adjustment. This occurred, presumably, because the observations are essentially judgments. This finding, which only came to light at this stage of follow-up measurement, underlines the vital importance of using the same set of observers for both before and after measures to control for inter-observer differences.

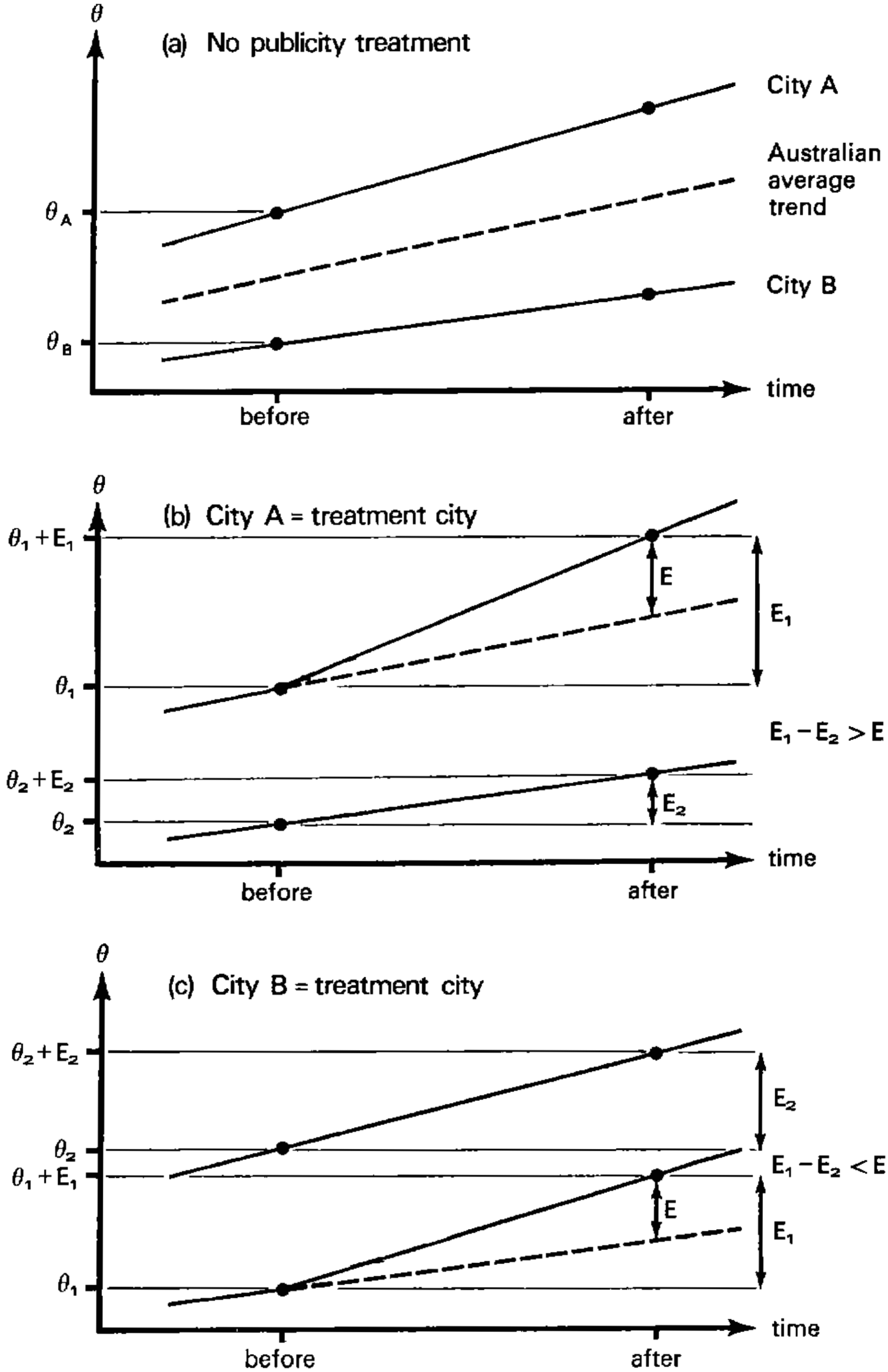
The data collected had, therefore, to be discarded and no insight was gained into the duration of the behavioural changes achieved.

Discussion

Clear evidence was found of significant behaviour change resulting from short, intense exposure to television publicity. However, the validity of the experiment depends on the adequacy of using disparate, unmatched areas to form the basis of experimental and control populations. Several reviewers strongly criticise this type of experimental design.[5, 15]

The objections are best described in terms of the hypothetical situation shown in Figure 3. Assume there was a general increasing trend throughout Australia in θ, the probability of a car occupant wearing his seat belt correctly and that city A's trend was increasing faster than the Australian average, while city B's trend was increasing more slowly. Figure 3 (a) illustrates these hypothetical cities at the time of the study, assuming no publicity occurred in either city. The probabilities θ_A and θ_B are shown as different for illustrative purposes, but this is not essential to the argument.

Figure 3 (b) shows the situation where city A was chosen as the treatment city and the publicity was presented between the 'before' and 'after' observation periods. Then $\theta_A = \theta_1$ and $\theta_B = \theta_2$, for comparison with Figure 2 (page 7). E is the true effect of the publicity (assumed positive), that is, E is the additional increase in the probability of correct wearing above the trend in city A. E_1 is the increase in θ_1 between the before and after periods in city A, and E_2 is the corresponding increase in θ_2 in city B. The method of analysis described in Appendix A tests whether $E_1 - E_2$ is different from zero. Note that, because of the different trends in the two cities, $E_1 - E_2$ is greater than E and hence the three-way chi-square test may conclude $E_1 - E_2$ is greater than zero when in fact E is zero.



θ
(a) No publicity treatment
City A
Australian average trend
City B
θ_A
θ_B
time
before
after
θ
(b) City A = treatment city
θ₁ + E₁
θ₁
θ₂ + E₂
θ₂
E
E₁
E₁ − E₂ > E
E₂
time
before
after
θ
(c) City B = treatment city
θ₂ + E₂
θ₂
θ₁ + E₁
θ₁
E₂
E₁ − E₂ < E
E
E₁
time
before
after

Figure 3

Consider now the reverse situation where city B is chosen as the treatment city, as shown in Figure 3 (c). In this case $E_1 - E_2$ is less than E and hence the three-way chi-square test may not reject the hypothesis that $E_1 - E_2$ equals zero when in fact E is greater than the minimum publicity effect it is desired to detect.

Clearly neither situation is satisfactory. The experimenter usually does not know which situation he is in, unless he monitors the trends in θ for some time before the experiment and is prepared to make the assumption that the trends would have held constant during the experiment. It was not expected that, in practice, changes in θ would be anything but small over the few weeks of the experiment nor that the trends would be greatly different from city to city. However, quite small changes in θ (due to the publicity) were considered possible and important to detect.

The solution adopted was to conduct an experiment with an arbitrary choice of treatment city (Experiment 1) followed by a further experiment in which the roles of treatment and control city are reversed (Experiment 2). The two experiments were conducted sufficiently far apart in time with the aim of avoiding any residual effect of the publicity in the first treatment city. By this procedure two estimates of E are obtained which may be compared and their average taken as a better estimate of the true effect of the publicity. Of course, the validity of this procedure relies upon the assumption that the relative slope of the trends in θ in the two cities does not change greatly over time; this is possible but unlikely. It was considered less likely than the probability of reaching a biased estimate of E if only one experiment were performed.

Experiment 2

Method

Experiment 1 was conducted in the period May to July 1973. Experiment 2 was conducted in November and December of the same year; it was considered unlikely that a residual effect of the short publicity program would have persisted for four months. It should be noted that the further measurements in Adelaide and Melbourne, which did not receive additional publicity, cannot be used to assess the long-term decay effects from Experiment 1 since the majority of observers were newly recruited and trained and inter-observer differences preclude comparisons.

All aspects of the general methodology and experimental procedure were identical to those employed in Experiment 1 except that the roles of the cities were reversed (Table 4).

Table 4: Experimental design

City	*Before measurements*	*Publicity*	*After measurements*
Adelaide . .	6 a.m.–Midnight 15–18 November same sites as for Experiment 1	CONTROL (Nil) NB formerly HIGH	6 a.m.–Midnight 6–9 December at identical sites
Melbourne .	As above . .	IMPURE CONTROL (Nil—although one channel had continued to screen the film irregularly since Experiment 1)	As above
Hobart . .	As above . .	HIGH (4 × 60 seconds per night on the sole commercial channel for 14 consecutive nights). NB formerly CONTROL	As above

Hobart, the smallest of the three cities, has only one commercial television channel, compared with three in Melbourne and Adelaide. To approximate the intensity achieved in the first experiment, the film was screened four times per night in 'prime' viewing time. In Adelaide, in Experiment 1, it had been screened three times per night on each of three channels. With 'random' channel selection it would theoretically have been possible to have observed the film nine times per night in Adelaide but this was considered extremely unlikely. In the absence of relevant viewing data it was considered that four times on the only available commercial channel in Hobart would be roughly equivalent.

Results

All occupants

Tables 5 and 6 summarise the results for all occupants. (The results are given in full in Appendix C.)

Table 5: Percent change × target behaviour × publicity condition

	Publicity condition		
	HIGH (*Hobart*)	*IMPURE CONTROL* (*Melbourne*)	*CONTROL* (*Adelaide*)
Per cent change in proportion 'tight' . .	15.4	−3.5	−0.2
Number of observations before/after . .	2583/2743	3701/3774	3101/3059
Per cent change in proportion buckle 'off hip' .	18.2	4.4	7.5
Number of observations before/after . .	2584/2736	3692/3773	3085/3038
Per cent change in proportion with 'no twist' .	8.2	1.1	4.4
Number of observations	2585/2743	3695/3773	3091/3060

Table 6: Significance levels of statistical tests

	High v. Control	*High v. Impure Control*	*Impure Control v. Control*
Tightness	< 0.001	< 0.001	N.S.
Buckle position . . .	< 0.001	< 0.001	N.S.
Twist	< 0.05	< 0.001	< 0.01

The High v. Impure Control differences were highly significant (< 0.001) for all three measures of correctness, as were the High v. Control differences except for the flatness measure which was significant at the 0.05 level. Melbourne appears to be equally as good a control city as Adelaide, even though it was a somewhat impure control. These results indicate a real effect of high levels of publicity (four screenings per night) on improving seat belt adjustment, confirming the effect found in Experiment 1. The percentage wearing their seat belts 'tight' or 'in-between' increased by 15.4 per cent from 48.2 per cent, 'no twist' by 8.2 per cent from 74.0 per cent, and buckle 'off-hip' by 18.2 per cent from 37.1 per cent. The magnitude of these changes are larger than observed in Adelaide in Experiment 1, suggesting perhaps that the publicity levels were not equally intense.

Controlled analyses

Only time of day, occupant age and day of week were investigated as potential 'controlling' descriptors for Experiment 2. In Hobart the apparent effect of the publicity on tightness was found to vary significantly with time of day (< 0.05) and on buckle position with time of day, age, and day of week (all < 0.05). The distribution of observations by time of day and day of week also varied significantly (< 0.01) between the before and after stages in Hobart. The control cities were not investigated. From these analyses it did not appear imperative to control for occupant age, but again a conservative stance was adopted and a decision made to control for time of day, day of week, and occupant age for all three measures of correctness.

The results from the controlled analyses are presented in full in Appendix C. It should be noted here that the results largely failed to confirm the differential publicity effects as a function of occupant age, time of day and day of week observed in Experiment 1.

Discussion

There can be no doubt that the television publicity demonstrating optimal seat belt adjustment increased the incidence of tightly adjusted belts, buckles located rearward of the hip and untwisted webbing. The successful replication of the results of the first experiment removes the possibility that the observed changes were due to factors other than the publicity program associated with the disparate areas used as experimental and control sites.

Successful replication through use of the reverse design also added confidence to the generality of the results. The publicity effect had been found in two capital cities which differ markedly in size, socio-economic composition, topography and road and traffic patterns.

The finding of no effect for a low level of publicity over a short duration was important for its implications on the level of expenditure required to effect behavioural change. Conversely, the intense level of publicity at which behaviour change occurred is very costly. Because of these program cost implications it was decided to conduct a third experiment varying both the intensity and duration of exposure to television publicity.

Experiment 3

Method

The third experiment was conducted during May and June 1974. The independent variables were publicity intensity (screenings per night) and duration (number of nights). Five cities were used, three State capitals and two major provincial cities in New South Wales.

All aspects of the general methodology and experimental procedure were identical to those employed in Experiments 1 and 2 with the following exceptions:

- observations ceased at 11 p.m. instead of midnight for operational reasons;
- measurements were taken in the control city three times instead of twice; two 'before' measurements to match the different starting dates for the experimental cities with long and short publicity duration, and one 'after' measurement.

The four publicity conditions were:

HIGH-SHORT (Newcastle)	—identical to that used in Hobart in Experiment 2 to provide a known base for comparison that is, four times per night for fourteen nights
HIGH-LONG (Wollongong)	—same intensity as above but twice the duration (twenty-eight nights)
LOW-SHORT (Brisbane)	—two-thirds of the intensity used in Adelaide in Experiment 1 but for the same duration; the low level used in Melbourne in Experiment 1, which produced no effect, was one-third of the Adelaide level
LOW-LONG (Perth)	—the intensity was only approximately equal to that for LOW-SHORT as Perth had only two commercial channels where Brisbane had three; the duration was twenty-eight nights

The experimental design is presented in Table 7.

Table 7

City	*Before measurements*	*Publicity*	*After measurements*
Wollongong .	6 a.m.–11 p.m. 2–5 May at a minimum of five sites	HIGH-LONG (4 × 60 seconds per night on the sole commercial channel for 28 consecutive nights).	6 a.m.–11 p.m. 6–9 June at identical sites
Newcastle	As above 16–19 May	HIGH-SHORT (4 × 60 seconds as above but for 14 nights).	As above
Perth . .	As above 2–5 May	LOW-LONG (2 × 60 seconds on each of the two commercial channels for 28 consecutive nights).	As above
Brisbane . .	As above 16–19 May	LOW-SHORT (2 × 60 seconds on each of the three commercial channels but for 14 nights).	As above
Sydney . .	As above 2–5 May and 16–19 May	CONTROL (Nil) . .	As above

Results

All occupants

In view of the number of treatments and the two sets of measurements in the control city the results are presented below separately by treatment. (The data are presented in full in Appendix D).

HIGH-SHORT: Table 8 presents the data in summary form.

Table 8: Per cent change × Target behaviour: HIGH-SHORT

	HIGH-SHORT (Newcastle)	*CONTROL (Sydney)*	*Result of significance test*
Percentage change in proportion 'tight'	7.6	4.6	<0.05
Number of observations Before/After	3043/3467	2514/2445	
Percentage change in prop. buckle 'off-hip'	12.9	8.5	<0.05
Number of observations Before/After	3028/3456	2495/2437	
Percentage change in prop. with 'no twist'	0.8	0.7	N.S.
Number of observations Before/After	3045/3462	2494/2438	

Since the treatment level was the same as that used in the first two experiments, the results provide conclusive confirmation of the effectiveness of this intensity and duration of television publicity in modifying behaviour. The effect, at least on the two behaviours known to be directly related to injury occurrence, has been observed in three disparate cities—Adelaide, Hobart and Newcastle.

HIGH-LONG: Table 9 contains the relevant data.

Table 9: Per cent change × target behaviour: HIGH-LONG

	HIGH-LONG (Wollongong)	*CONTROL (Sydney)*	*Result of significance test*
Percentage change in proportion 'tight'	9.9	7.7	
Number of observations Before/After	2511/2214	2895/2445	N.S.
Percentage change in proportion buckle 'off-hip'	11.5	11.4	
Number of observations Before/After	2499/2211	2884/2437	N.S.
Percentage change in proportion with 'no twist'	4.8	0.2	<0.01
Number of observations Before/After	2504/2214	2875/2438	

Only one of the target behaviours—the absence of webbing twist—is there evidence of a significant increase in incidence compared with the control city. This is somewhat surprising since intensity was the same as that shown to be effective in three experiments and the duration was longer; the logical expectation would have been for a stronger effect.

It is interesting to note that the management of the sole commercial channel in Wollongong wrote to the authors criticising the use of only one television film in a saturation campaign and pointed out that they had received several complaints from viewers that its screening was too frequent. It is possible to speculate, that the extended duration of high frequency screening led to a cancelling effect. Whether this treatment would have succeeded had a variety of films with the same message been used is unknown and should be subjected to study. It is only possible to conclude that, with a single communication, doubling the duration of exposure does not increase effectiveness.

LOW-LONG: Table 10 summarises the results.

Table 10: Per cent change × target behaviour: LOW-LONG

	LOW-LONG (Perth)	*CONTROL (Sydney)*	*Result of significance test*
Percentage change in proportion 'tight'	7.3	7.7	
Number of observations Before/After	3732/4101	2895/2445	N.S.
Percentage change in proportion buckle 'off-hip'	12.0	11.4	
Number of observations Before/After	3728/4100	2884/2437	N.S.
Percentage change in proportion with 'no twist'	4.8	0.2	
Number of observations Before/After	3731/4101	2875/2433	< 0.01

It is clear from Table 10 that the LOW-LONG publicity condition had little effect on the incidence of the target behaviours, with the exception of increasing the proportion of vehicle occupants who wore their belts without twist in the webbing.

LOW-SHORT: The summary results are contained in Table 11.

Table 11: Per cent change × target behaviour: LOW-SHORT

	LOW-SHORT (Brisbane)	*CONTROL (Sydney)*	*Result of significance test*
Percentage change in proportion 'tight'	11.9	4.6	
Number of observations Before/After	2607/2819	2514/2445	< 0.001
Percentage change in proportion buckle 'off-hip'	7.6	8.5	
Number of observations Before/After	2600/2818	2495/2437	N.S.
Percentage change in proportion with 'no twist'	3.4	0.7	
Number of observations Before/After	2603/2819	2494/2438	N.S.

Only for the target behaviour of tightness of adjustment did the LOW-SHORT publicity condition have a significant effect. It is interesting to note that both of the low intensity publicity conditions produced change in one of the three target behaviours. The intensity level, as pointed out earlier, was roughly mid-way between the 'low' level used in Melbourne, which produced no effect, and the 'high' level used in Adelaide in Experiment 1.

In summary, the HIGH-SHORT publicity condition produced the most consistent effect, resulting in positive changes to two of the three target behaviours. In contrast the HIGH-LONG condition, which would have passed through the high-short phase, appears to have had a cancelling effect; the two criteria significant under the HIGH-SHORT condition are not significant under HIGH-LONG.

The LOW-LONG condition produced a similar change to webbing twist as did HIGH-LONG but apparently zero change on the other two target behaviours. Again there appeared to be a detrimental effect of the LOW-LONG condition compared with LOW-SHORT, at least as far as tightness is concerned. The LOW-SHORT condition resulted in a positive change to tightness, which was not apparent for LOW-LONG.

The most effective publicity treatment is clearly one combining saturation intensity with short duration. Increasing the duration, while still using a single communication, appeared to dampen effectiveness. Decreasing intensity while maintaining a short duration, with its implications for program cost reduction, did not result in behavioural change for all target behaviours. Similarly, decreasing intensity but doubling duration was not effective.

Controlled analyses

In Experiment 3, the five descriptors of each occupant observed were augmented by another two:

- wearing overcoat (yes/no);
- weather.

These were added since the measurements were taken in winter and the range of climates across cities was considerable.

The results from the controlled analyses are given in full in Appendix D.

Summary and conclusions

Summary of results

Table 12 summarises the outcome of the statistical tests conducted on the data for all front seat occupants in each of the three experiments.

Table 12: Summary of results of significance tests (all occupants)

Publicity condition	*Adjustment criteria* *Tightness*	*Buckle position*	*Webbing twist*
Experiment 1			
High *v.* Control	**	***	***
Low *v.* Control	—	—	—
High *v.* Low	***	***	***
Experiment 2			
High *v.* Control	***	***	*
High *v.* Impure Control	***	***	***
Control *v.* Impure Control	—	—	**
Experiment 3			
High-Short *v.* Control	*	*	—
High-Long *v.* Control	—	—	**
Low-Short *v.* Control	***	—	—
Low-Long *v.* Control	—	—	**

Key * improvement significant at $p < 0.05$
** improvement significant at $p < 0.01$
*** improvement significant at $p < 0.001$
— no significant difference

In Experiment 1, significant increases in the incidence of each of the three target behaviours were observed in the city receiving short, intense exposure to television publicity. The exposure comprised three screenings of a sixty second film on each commercial channel for fourteen consecutive evenings in the prime viewing time of 6 p.m. to 9 p.m.

To control for the possibility that the behavioural changes were due to factors other than the publicity—the experimental and control groups were drawn from disparate cities—a second experiment was conducted in which the same cities were used but their roles were reversed. The results were replicated, providing confirmation that the publicity treatment was responsible for the observed behavioural change.

In Experiment 3, a range of publicity treatments was examined, intensity and duration of exposure being the independent variables. Short, intense exposure, similar to that used in the first two experiments, was the only treatment to provide consistent results; the incidence of two of the three target behaviours significantly increased. For the other three publicity treatments behavioural change was observed for only one of the target behaviours.

Importance of the behavioural changes

The ultimate goal of the publicity program was to reduce the incidence of those injuries to vehicle occupants which arise from sub-optimum seat belt adjustment. For two reasons, changes in injury patterns could not be used as the dependent variable; sufficiently detailed injury data are not routinely available and seat belt wearing in crashes is not reliably reported. It was not possible therefore to undertake cost-benefit calculations.

Although the target behaviours were chosen because of their demonstrated role in injury causation it is difficult to estimate the practical importance of increases in their incidence, particularly when the incidence of optimal adjustment among seat belt wearers remains substantially less than 100 per cent. In other words, although statistically significant behavioural change has been achieved, the magnitude of the benefit in terms of injury reduction remains unknown.

Table 13 shows the observed base levels in each of the cities receiving the short intense publicity treatment and the proportionate changes achieved. While these figures provide some guide to the magnitude of the behavioural changes caution is necessary since, strictly, each requires interpretation in terms of the concomitant change in the relevant control population.

Table 13: Magnitude of behavioural change (all occupants)

	City receiving short, intense publicity								
	Adelaide (experiment 1)			*Hobart (experiment 2)*			*Newcastle (experiment 3)*		
Target behaviour	*Before*	*After*	*Prop. change*	*Before*	*After*	*Prop. change*	*Before*	*After*	*Prop. change*
per cent 'tight'	61	70	15	48	63	31	68	76	12
per cent buckle 'off-hip'	53	65	23	37	55	49	41	54	32
per cent 'no twist'	75	81	8	74	82	11	80	81	1*

* the only statistically insignificant increase (compared with relevant control).

The smallest proportionate increases were for the target behaviour of webbing twist, in part no doubt because the scope for change was small—the base levels of desired behaviour ranged between 74 and 80 per cent. The largest proportionate increases were for the target behaviour of buckle position, ranging from 23 per cent to 49 per cent. Again the magnitude of the change appeared to depend, in part, on the magnitude of the base level of the desired behaviour.

Although no definitive answer can be given to the question of the practical importance of the observed behavioural changes it is clear that the changes, in addition to being statistically significant, are, by and large, substantial.

Implications for publicity as a countermeasure

It is very difficult to generalise from the results of one type of publicity program to others. However, given consistent results in three disparate cities at three points in time—the first and third experiments being twelve months apart—a degree of generalisation seems warranted.

The features of the program considered to have contributed to its success, and hence to be potentially applicable to other publicity programs, are discussed below.

Target behaviours. The target behaviours were:
- directly related to injury causation;
- specific and capable of unambiguous demonstration;
- aspects of a widespread parent behaviour (seat belt wearing) towards which the target group had favourable attitudes.

Nature of communication. The communication:
- used a message source of high credibility and high status;
- used a mild fear setting as a motivating factor;
- comprised a specific, concrete demonstration of the target behaviours.

Generalisation in this area of creative approach is hazardous. Given the successful behavioural changes achieved it would be advantageous to conduct a further experiment using nature of the communication as the independent variable.

Intensity and duration of exposure. Short, intense exposure, as defined operationally in these experiments, seems an essential ingredient. Even here, however, there is an interplay with the nature of the communication. In Experiment 3 anecdotal evidence was related which indicated that the 'long', intense publicity treatment failed in part because there was no variety in the message, only one film having been used.

None of the features discussed above is new to workers in the field of communication. Indeed, most have been put forward as guidelines for the development of road safety publicity programs.[2, 11, 12, 16] What this series of experiments has done is reinforce the utility of these communication principles and demonstrate the benefits of their use as a basis for program design and implementation.

Conclusions

Three major conclusions can be drawn from this series of experiments. First, and most importantly, television publicity has been shown to be a practical countermeasure to the immediate problem of sub-optimal seat belt adjustment and a valuable adjunct to the longer term solution provided by a re-design of the seat belt system. Until the effects of the re-design have spread through the vehicle population, repetitions of the publicity program in short, intense bursts are warranted and can be undertaken with confident expectation of favourable behaviour change.

Secondly, the validity and utility of a limited range of communication principles has been reinforced. These principles should be utilised whenever possible in campaign design and implication.

Thirdly, the series of experiments has demonstrated that research to measure the actual effects in the 'real world' of a full-scale publicity program can be successfully undertaken.

References

1. Andreassend, D. C., 'The Effects of Compulsory Seat Belt Wearing Legislation in Victoria', *National Road Safety Symposium, Canberra, 14–16 March 1972*. Australian Government Publishing Service, Canberra, 1972.
2. Avery, G. C., 'The Role of Communication and Propaganda in Traffic Safety', *Report* 3/73 Traffic Accident Research Unit, Department of Motor Transport, Sydney, 1973.
3. Cameron, M. H., Accident Risks—Concept and Analysis, M.Sc. Thesis, University of Melbourne, 1969.
4. Freedman, K., Champion, P. and Henderson, M. Seat Belts: A Survey of Usage and Attitudes. *Report* 2/71 Traffic Accident Research Unit, Department of Motor Transport, Sydney, 1971.
5. Haskins, J. B., 'Evaluative Research on the Effects of Mass Communication Safety Campaigns: a Methodological Critique', *Journal of Safety Research*, 1970, 2(2), 86–96.
6. Herbert, D. C., Vazey, B. A. and Henderson, J. M. ,'Seat Belt Crash Performance in Australia', *Proceedings of Seat Belt Seminar*, Melbourne, 9–11 March, 1976. Australian Government Publishing Service, Canberra, 1977.
7. Holt, B. W. Occupant Injury from Seat Belt Use. Paper presented to the Road Trauma Section of the 48th General Scientific Meeting of the Royal Australasian College of Surgeons, New Zealand, 15–22 August, 1975.
8. Kendall, M. G. and Stuart, A. *The Advanced Theory of Statistics, Volume 2, Inference and Relationship*, Griffin, London, 1973.
9. Milne, P. W., *The Fitting and Wearing of Seat Belts in Australia: The History of a Successful Countermeasure*, Office of Road Safety Report OR1, Department of Transport, Melbourne, 1979.
10. Morris, J. P., *Road Safety Publicity—Quantifying the Effectiveness of Public Service Advertising*, The Advertising Association, London, April 1972.
11. Organisation for Economic Co-operation and Development. *Road Safety Campaigns: Design and Evaluation* OECD, Paris, 1971.
12. Organisation for Economic Co-operation and Development. *Manual on Road Safety Campaigns: Guidelines for their Preparation, Design and Effectiveness Control* OECD, Paris, 1975.
13. Robertson, L. S., Kelley, A. B., O'Neill, B., Wiscan, W., Eiswirth, R. S. and Haddon, W. (Jr), A Controlled Study of the Effect of Television Messages on Safety Belt Use. Insurance Institute for Highway Safety, Washington, D.C., June 1972.
14. Ryan, G. A. and Baldwin, R. J., In-Depth Study of Seat Belted Accidents. Report to the Department of Shipping and Transport, Melbourne, December 1972.
15. Spolander, K., Evaluation of Safety Campaigns in Terms of Behavioural Change. Paper presented at the International Conference on the Design of Safety Campaigns, Italian Ministry of Public Works, Rome, 13–15 October, 1971.
16. Wilde, G. J. S., Road Safety Publicity: Status Quo and Future Developments. Paper presented at the International Conference on the Design of Safety Campaigns, Italian Ministry of Public Works, Rome, 13–15 October, 1971.

APPENDIX A

Statistical considerations

Method of analysis

The method of analysis was the three-way chi-square test for second-order interaction in complex contingency tables for three categorical variables.[8] The three categorical variables in this study were:

- city (treatment *v.* control)
- time (before *v.* after publicity campaign)
- measure of correctness (K levels)

Let i, j and k be subscripts denoting the particular levels of these three categorical variables, respectively. Subscripts i and j each only take values 1 and 2, but let their maximum values be denoted by I and J, respectively, for generality, i.e. i takes values 1, 2, . . ., I and j takes values 1, 2, . . ., J.

The measure of correctness variable took on three different roles in turn, namely

- tightness (three levels)
- flatness (two levels)
- buckle position (three levels)

These three measures each have a nominally 'correct' level describing belt wearing, and the probabilities in Figure 2 (Page 7) can be thought of as the probability of an occupant wearing his belt 'correctly' on any particular measure. Another measure was conceived during the design of this study, namely 'correct on all three criteria' *v.* 'incorrect on any one or more'. This measure was not explicitly used in the analysis, but was used as the basis for calculating the required sample size (see next section).

Let n_{ij} be the number of car occupants sampled in the i-th city during the j-th time period. Let N be the sum of the n_{ij}, summed over i and j. Let Π_{ijk} be the probability that any one of the N sampled occupants was observed in the i-th city during the j-th time period wearing his seat belt at the k-th level of correctness. Then a suitable null hypothesis equivalent to 'no effect of the publicity' is:

H_0: $\Pi_{ijk} = \Pi_{ij.}\ \Pi_{i.k}\ \Pi_{.jk}/(\Pi_{i..}\ \Pi_{.j.}\ \Pi_{..k})$, all i, j, k,

where a dot subscript represents summation over the subscript it replaces, for instance

$\Pi_{ij.} = \sum_k \Pi_{ijk},\ \Pi_{i..} = \sum_j \Pi_{ij.} = \sum_{j,k} \Pi_{ijk},$

and, of course, $\sum_i \Pi_{i..} = 1$.

That H_0 is equivalent to no effect of the publicity is demonstrated as follows. $\Pi_{ij.}$ is the probability that an occupant was observed in the i-th city during the j-th time period (and hence equals n_{ij}/N), so $\Pi_{ijk}/\Pi_{ij.}$ is simply the probability that any one occupant observed in the (i, j) city/time was wearing his seat belt at the k-th level of correctness. Now if H_0 is true, then:

$$\begin{aligned}\Pi_{ijk}/\Pi_{ij.} &= \Pi_{i.k}\ \Pi_{jk.}/(\Pi_{i..}\ \Pi_{.j.}\ \Pi_{..k})\\ &= \alpha_{ik}\ .\ \beta_{jk},\ \text{for some } \alpha_{ik},\ \beta_{jk},\end{aligned}$$

that is, the probability of wearing at the k-th level of correctness is the product of a number specific to the city (α_{ik}) and a number specific to the time period (β_{jk}), for each k. If the publicity had an effect, then clearly $\Pi_{ijk}/\Pi_{ij.}$ could not be expressed as such a product, for some k. Hence H_0 is equivalent to no effect of the publicity. The hypothesis H_0 is also known as the hypothesis of no second-order interaction between the three categorical variables.

It remains to give a test for H_0 against the general alternative hypothesis H_1, the falsity of H_0. Let n_{ijk} be the number of occupants observed in the (i, j) city/time wearing their seat belts at the k-th level of correctness. Using the dot subscript convention, it follows that:

$$n_{ij.} = \sum_k n_{ijk} = n_{ij}, \text{ all } i, j.$$

By an argument similar to Cameron[3] it follows that, when H_0 is true,

$$X^2 = \sum_{ijk} \frac{(n_{ijk} - N n_{ij.}\, n_{i.k}\, n_{.jk}/(n_{i..}\, n_{.j.}\, n_{..k}))^2}{N n_{ij.}\, n_{i.k}\, n_{.jk}/(n_{i..}\, n_{.j.} n_{..k})}$$

is asymptotically (as N tends to infinity) distributed like a chi-square variable on $(I - 1)(J - 1)(K - 1)$ degrees of freedom. This criterion was compared with the 100α per cent critical values of a chi-square distribution with appropriate degrees of freedom. Values of X^2 exceeding the 5 per cent value were deemed significant, and values exceeding higher critical values (that is lower α) were especially noted. If X^2 exceeded the 100α per cent critical value, then either

- a rare event has occurred (with probability α); or
- H_0 is false.

It is usual to conclude the latter. Then α is the probability of a Type I error, that is reject H_0 when it is true. When a large number of such tests are made (as were in this study), it should be expected that a proportion α of the conclusions that H_0 is false (that is a publicity effect exists) are Type I errors. This proportion was 5 per cent in this study, but sufficient information is given with the results for the reader to modify the choice of α and hence the proportion of Type I errors.

A rather large probability of Type I error was chosen to give the three-way chi-square test maximum power to find a publicity effect if it existed. The power of a statistical test is the probability of rejecting H_0 when it is false. A Type II error occurs if H_0 is accepted when it is false, and this has probability of one minus the power. In this study Type II errors were considered more important to avoid than Type I errors. A Type I error might mean that public funds will be wasted in future by showing an ineffective television film, but ineffective publicity is already common in the road safety field. However a Type II error would mean that knowledge of a truly effective form of publicity has been missed, along with all the advantages of further presentations of this particular television film plus guidelines for the production of further publicity campaigns in the road safety field.

The probability of a Type II error can be made arbitrarily small by appropriate choice of the total sample size, N. This is the subject of the following section.

Sample size

It will be shown that the power of the three-way chi-square test depends on the total sample size (in both cities and time periods), the publicity effect, and α, the probability of a Type I error. It follows that the sample size used depended on the likely publicity effect and the probability (power) with which the investigators wished to detect this effect.

If H_1 is true (that is a publicity effect exists), then X^2 has asymptotically (as N tends to infinity) the non-central chi-square distribution on $\nu = (I-1)(J-1)(K-1)$ degrees of freedom and non-centrality parameter γ, and with distribution function $\chi^2(\nu, \gamma)$. Let $\Pi_{ijk}^{(0)}$ and $\Pi_{ijk}^{(1)}$ be the hypothetical values of Π_{ijk} in H_0 and H_1, respectively; that is:

$$\Pi_{ijk}^{(0)} = \Pi_{ij.}\,\Pi_{i.k}\,\Pi_{.jk}/(\Pi_{i..}\,\Pi_{.j.}\,\Pi_{..k})$$

but $\Pi_{ijk}^{(1)}$ is undefined as yet. Then, according to Kendall and Stuart[8] the non-centrality parameter is:

$$\gamma = N \sum_{ijk} (\Pi_{ijk}^{(1)} - \Pi_{ijk}^{(0)})^2/\Pi_{ijk}^{(0)} = N\kappa^2, \text{ say.}$$

κ^2 depends on the effect of the publicity. It is non-negative and takes value zero only when $\Pi_{ijk}^{(1)}$ equals $\Pi_{ijk}^{(0)}$ for all i, j and k, that is when the publicity effect is zero.

For a particular critical value $\chi_a^2(\nu)$ of the ordinary (central) chi-square distribution, the probability of a Type II error when H_1 is true is:

$$\beta = \text{Prob}\,(X^2 < \chi_a^2\ (\nu)) \approx \int_0^{\chi_a^2(\nu)} d\chi^2(\nu, \gamma)$$

for large N. Hence the power of the test is approximately:

$$P = 1 - \beta = \int_{\chi_a^2}^{\infty} (\nu)\, d\,\chi^2(\nu, \gamma)$$

It can be seen that P, κ^2, N and a are all intimately related. For fixed a and κ^2, it also can be seen that P can be made as close to 1 as desired by increasing N, from Kendall and Stuart.[8] The intention in this study was to choose N such that the power would be close to 1 for the minimum, non-trivial value of κ^2 which the publicity could cause. Then the power would be even greater for larger publicity effects, but perhaps considerably smaller for smaller publicity effects.

In the case where the measure of correctness is a dichotomy (that is $K = 2$), X^2 has a chi-square distribution on one degree of freedom. In this situation only, the power P can be re-written as:

$$P = \Phi\,(\sqrt{\chi_a^2(1)} - \sqrt{\gamma}) + \Phi\,(\sqrt{\chi_a^2(1)} + \sqrt{\gamma})$$

where

$$\Phi(x) = \int_x^{\infty} (2\Pi)^{-\frac{1}{2}} \exp\,(-\tfrac{1}{2}\,x^2)\,dx,$$

the normal distribution cumulative density function. An approximation for the power function of chi-square test on more than one degree of freedom also exists but does not lead to an analytical solution for γ^3. Accordingly, it was decided to base the calculation of N on the dichotomous, overall measure of correctness, that is 'correct on all three criteria' v. 'incorrect on any one or more'.

It remained to express κ^2 in terms of E_1 and E_2 (Figure 2, page 7) and hence to calculate N for some minimum likely publicity effect measured by, say, $E_1 - E_2$.

However the mathematics of this proved intractible except under certain, perhaps unrealistic, assumptions as follows:

- $n_{ij} = \frac{1}{2}n_{i.}$, all i, j
- $\theta_1 = \theta_2 = \theta$
- $E_2 \approx o$ and E_1 small relative to θ

The result is then:
$\kappa^2 = (E_1 - E_2)^2/[4(\theta + E_1)(1 - \theta - E_1)]$,
approximately.

It was known from other studies that the proportion of car occupants with belts buckled who wore their belt correctly on all three criteria was of the order of 20 per cent, approximately (that is $\theta = 0.2$).[1,14] It was also known from other studies that the effect of the particular type of publicity under evaluation was likely to be small.[11] Changes in the base level of 20 per cent correct to a new level of 23 per cent or 24 per cent were minimum changes which might have been expected (that is $E_1 = 0.03$ or 0.04). Changes in the probability of correct wearing in the control city were assumed to be small during a short publicity campaign, so E_2 was taken a zero for the sample size calculation. It was decided to take $E_1 - E_2 = 0.035$ as the minimum, non-trivial publicity effect to be detected with high power (at least 0.95).

The final step was to calculate κ^2 for $E_1 - E_2$ and E_1 both equal 0.035 and $\theta = 0.2$, γ for $\alpha = 0.05$ and some P exceeding 0.95, and hence $N = \gamma/\kappa^2$. However the assumptions under which κ^2 was derived were still questionable, particularly the first two. The first assumption says that the number of observations in the before and after periods were identical, for both the treatment and control cities separately. While this was intended to be the case approximately by scheduling the same observation times in the before and after periods, it was acknowledged that variation in traffic patterns and weather would be such that exact equality was most unlikely. The second assumption says that the base probabilities of correct seat belt wearing were the same in the treatment and control cities. This was known to be only approximately the case. The key point was that, while all three assumptions are of the type that would normally be made when substituting empirical values in an equation for κ^2, the validity of the κ^2 equation itself should not depend on these assumptions which were strictly unrealistic in this study.

Another, perhaps more important, consideration affected the choice of an alternative method to calculate the sample size actually used. This was the fear that the control city would in fact turn out to be an 'impure' control, due to some major, unplanned environmental change which could be suspected to affect the correctness of seat belt wearing in that city alone. This would destroy the experimental design used but it would be possible to salvage a (much less satisfactory) 'before and after' study from observations made in the treatment city. Accordingly, it was decided that the sample size in the treatment city should be such that it would be capable of detecting $E_1 = 0.035$ with high power in a simple 'before and after' study. The method for calculating the sample size in this case is given in the next section, the exhaustive dichotomy case being relevant. For $\theta = 0.2$ and equal probability of an observation in the before and after periods, a total sample of 8000 will detect $E_1 = 0.035$ with power 0.9667 when $\alpha = 0.05$. Since a total sample of 8000 (that is 4000 in each of the before and after periods) in each city was compatible with the budgeting and observation scheduling restrictions, this was taken as the sample size.

It remained to check that $N = 2 \times 8000 = 16\,000$ would be sufficient observations to detect a publicity effect $E_1 - E_2$ of at least 0.035 with high power when the three-way chi-square test was used. It was necessary to use the κ^2 equation derived under suspect assumptions. It turned out that $N = 16\,000$ will detect a publicity effect $E_1 - E_2 = 0.025$ with power 0.9667 when $\alpha = 0.05$. Larger publicity effects would, of course, be detected with higher power. That is, observations on 4000 occupants in each city in each time period would be sufficient to detect a change from 20 per cent correct to 22½ per cent or more correct in the treatment city, if no change is assumed in the control city.

It was also known that about 20 per cent of drivers and passengers did not wear seat belts even when fitted, at the time of the study (after compulsory seat belt wearing legislation in all Australian States and Territories). Hence it was decided that about 5000 subjects with seat belts fitted to their seating positions would need to be sampled in each period in each city to achieve the required total sample size. This determined the number of observation sites required in each city.

Sample sizes for 2 × 2 chi-square tests of before-and-after experiments

There are basically two types of before-and-after experiments, depending on the response variable, namely:

- exhaustive dichotomy (Z *v.* not Z), or
- treatment *v.* control (T *v.* C).

The method of choosing sample sizes does not depend on the type of experiment, but the final choice of parameters will. For generality define the response variable as R *v.* R¹. Then the table of observed frequencies will appear as follows:

	Before	After	
R	b	a	$b + a$
R^1	B	A	$B + A$
	$b + B$	$a + A$	N

In the before period, let θ be the probability of an observation having response R. This is sometimes known *a priori* or can be estimated by $b/(b + B)$ if the before observations have already been made. If we wish to test for a change in θ in the after period, the appropriate statistic is

$$X^2 = \frac{N(\,|\,aB - bA\,|\, - \tfrac{1}{2}N)^2}{(b + a)(B + A)(b + B)(a + A)}$$

which has approximately a chi-square distribution on 1 degree of freedom. X^2 incorporates Yates' correction for continuity.

X^2 can be used to test the null hypothesis

H_0: no change in θ

against the general alternative hypothesis

H_1: θ increases by E.

When H_0 is true, X^2 has the ordinary (central) chi-square distribution. When H_1 is true, X^2 has the non-central chi-square distribution with non-centrality parameter γ, where $\gamma = NK^2$. If Π_{ij} is the probability of an observation falling in the (i, j) cell in the above table, then

$$K^2 = \sum_{ij} (\Pi_{ij}^{(1)} - \Pi_{ij}^{(0)})^2 / \Pi_{ij}^{(0)},$$

where $\Pi_{ij}^{(0)}$ and $\Pi_{ij}^{(1)}$ are the hypothetical values of Π_{ij} under the hypotheses H_0 and H_1 respectively. Now H_0 is equivalent to

H_0: $\Pi_{ij} = \Pi_{i.} \, \Pi_{.j}$, for all i, j

where

$$\Pi_{i.} = \sum_j \Pi_{ij} \text{ and } \Pi_{.j} = \sum_i \Pi_{ij}, \text{ so } \Pi_{ij}^{(0)} = \Pi_{i.} \, \Pi_{.j}.$$

However the $\Pi_{ij}^{(1)}$ depend on E, as will be seen later.

If X^2 does not exceed χ_α^2 (1) when H_1 is true, then this is a Type II error, that is not reject H_0 when it is false. The probability of this is denoted β, and $P = 1 - \beta$ is the power of the test, that is the probability of rejecting H_0 when H_1 is true. P depends on α and γ, and γ in turn depends on N and E. If we want the test to reject H_0 when a particular non-zero value of E occurs, we must collect sufficient observations N such that the power P is close to 1.

If p and q such that $p + q = 1$ were the probabilities of an observation in the before and after periods respectively (p and q depend on the length of the observation periods and are normally under control of the researcher), then the $\Pi_{ij}^{(1)}$ are as in the following table, where $r = q/p$ and $F = 1 + E/\theta$:

	Before	After	
R	θp	$\theta q \, . \, F$	$\theta p(1 + rF)$
R^1	$(1 - \theta)p$	$(1 - \theta F)q$	$1 - \theta p(1 + rF)$
	p	q	1

It follows that

$$K^2 = \frac{q\theta(1 - F)^2}{(1 + rF)\,(1 - \theta p(1 + rF))},$$

Now for a chi-square test on 1 degree of freedom only

$$P = \Phi(\sqrt{\chi_\alpha^2(1)} - \sqrt{\gamma}) + \Phi(\sqrt{\chi_\alpha^2(1)} + \sqrt{\gamma})$$

where

$$\Phi(x) = \int_x^\infty (2\Pi)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}x^2)\, dx,$$

the normal distribution cumulative density function. For any reasonable γ,

$$\Phi(\sqrt{\chi_{\alpha}^{2}(1)} + \sqrt{\gamma}) \approx 0,$$

and hence

$$P \approx \Phi(\sqrt{\chi_{\alpha}^{2}(1)} - \sqrt{\gamma}).$$

Therefore

$$\gamma \approx (\sqrt{\chi_{\alpha}^{2}(1)} - \Phi^{-1}(P))^{2}, \text{ where } \Phi^{-1} \text{ is the inverse function of } \Phi$$

$$= (\Phi^{-1}(\tfrac{1}{2}\alpha) + \Phi^{-1}(1 - P))^{2}.$$

Finally,

$$N = \gamma/\kappa^{2} \approx \frac{(\Phi^{-1}(\frac{1}{2}\alpha) + \Phi^{-1}(1 - P))^{2}(1 + rF)(1 - \theta p(1 + rF))}{q\theta(1 - F)^{2}}$$

Frequently used values of Φ^{-1} are tabulated below:

x	$\Phi^{-1}(x)$
0.05	1.645
0.025	1.960
0.01	2.326
0.005	2.576
0.001	3.090
0.0005	3.291

Consider the two different types of response variable in turn.

1. *Exhaustive dichotomy*

This is the simpler of the two cases since there is no essential relationship between the parameters θ, p and q. It is usual, though not essential, to have equi-length before and after periods, in which case $p = q = \frac{1}{2}$ approximately. Then $r = 1$ and it follows that

$$N = \frac{(\Phi^{-1}(\frac{1}{2}\alpha) + \Phi^{-1}(1 - P))^{2}(1 + F)(2 - \theta(1 + F))}{\theta(1 - F)^{2}}.$$

2. *Treatment v. control*

For equi-length before and after periods, it would be usual to expect an observation from the control group to occur with equal probability both before and after. This requires

$$(1 - \theta)p = (1 - \theta F)q$$

and hence that

$$p = (1 - \theta F)/(2 - \theta(1 + F)).$$

It follows that

$$K^{2} = \frac{\theta(1 - F)^{2}}{2(1 + F - 2\theta F)}$$

and hence that

$$N = \frac{2(\Phi^{-1}(\frac{1}{2}\alpha) + \Phi^{-1}(1 - P))^{2}(1 + F - 2\theta F)}{\theta(1 - F)^{2}}$$

APPENDIX B

Detailed results—Experiment 1

Need for controlled analysis

Only the Adelaide data were investigated in detail to determine if there was need for controlled analysis. The apparent effect of the publicity on tightness was found to vary significantly with time of day (< 0.001) and age (< 0.05), and on buckle position with the time of day (< 0.01). The relative distribution of observations by time of day and age also varied significantly between the before and after stages in Adelaide. No significant variation of the apparent publicity effect by day of week was found, but both Melbourne and Hobart had significant changes in the distribution of observations by day of week. Melbourne was not investigated in detail so it was not known if it was necessary to control for day of week. Accordingly, it was decided to adopt a conservative stance and control for time of day, age and day of week separately when analysing the effect of the publicity on each of the three measures of correctness.

The results from the controlled analyses are shown in Tables I and II, together with the results for all front occupants for comparison.

Summary of controlled analyses

A problem with summarising controlled analyses was that although the need for a controlled analysis depended on the apparent effect of the publicity varying by occupant type in the treatment city, the actual effect as measured by the three-dimensional chi-square test for each occupant type separately need not vary, nor vary in the same way.

In summarising the results it was decided to note particularly those controlled analyses for which the actual effect of the publicity on the specific occupant type was quite different from that on all occupants. This exception reporting method was not entirely satisfactory because, being based on subsets of the data, the three-dimensional chi-square tests were less powerful than for all occupants. Hence there was less chance of detecting a real publicity effect of a specific size among particular occupant types than for all front occupants.

Also, because subsets of the data were used in the controlled analyses, extreme changes in each measure of correctness of car occupants in the control city occurred more frequently, where little or no change was expected. If such an extreme change *in itself* led to an apparent exceptional publicity effect on particular occupants, then this type of result was discounted as being evidence of an exceptional, true publicity effect on those occupants.

The results from the controlled analyses were summarised for each treatment city in turn, reporting particular results which appear to be strong exceptions to the overall publicity effects found for all front occupants.

Adelaide (HIGH): The following results are exceptions to the effects of a high level of publicity on:

- 'tight' (change relative to control 5.9 per cent for all occupants)
  - age 50 and over (+ 17 per cent relative change)
  - Thursday (+ 19 per cent)
  - 6.00 to 8.59 a.m. (+ 20 per cent)
  - Saturday (– 4 per cent)
- buckle 'off-hip' (relative change 9.5 per cent for all occupants)
  - age 50 and over (+ 17 per cent)
  - Thursday (+ 18 per cent)
  - 6.00 to 8.59 a.m. (+ 26 per cent)
  - 9.00 to 11.59 p.m. (+ 22 per cent)
- 'no twist' (relative change 6.4 per cent for all occupants)
  - age 50 and over (+ 9 per cent)
  - Thursday (+ 19 per cent)
  - 9.00 to 11.59 p.m. (+ 11 per cent)
  - Saturday (0 per cent)

Melbourne (LOW): The following results are exceptions to the effects of a low level of publicity on:

- 'tight' (relative change 0.2 per cent for all occupants)
  - age 50 and over (+ 11 per cent)
  - 3.00 to 5.59 p.m. (+ 11 per cent)
- buckle 'off hip' (relative change 0.1 per cent for all occupants)
  - age 50 and over (+ 10 per cent)
  - Friday (– 10 per cent)
  - 3.00 to 5.59 p.m. (+ 7 per cent)
- 'no twist' (relative change 1.9 per cent for all occupants)
  - 3.00 to 5.59 p.m. (+ 9 per cent)

Table I: Results of Experiment 1 controlled for age and day of week

Experiment 1	*All occupants*			*Age 8—29*			*Age 30—49*			*Age 50+*			*Thursday*			*Friday*			*Saturday*			*Sunday*		
City	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel*	*Melb.*	*Hob.*
TIGHTNESS																								
Tight or In-between—																								
Base level before (%)	61 0	57 1	67.2	65	57	68	64	59	66	49	52	68	55	58	66	56	60	64	65	59	66	66	52	74
Change After-Before	8.7	3.0	2.8	6	4	4	8	1	5	14	8	−3	19	6	0	12	3	4	2	3	6	4	0	−1
Significance / x^2 (1df)—																								
Adelaide *v.* Hobart	H		10.4	.		0.7	.		1.2	V		16.5	V		24.3	S		4.1	.		1.6	.		2.2
Melbourne *v.* Hobart	.		0 2	.		0.1	.		2.6	S		6.0	.		2.3	.		0.2	.		1 2	.		0.4
Adelaide *v.* Melbourne	V		14.0	.		1 5	H		9 6	.		1.8	V		14.7	H		7.3	.		0.1	.		2.9
BUCKLE POSITION																								
Off-hip																								
Base Level Before (%)	52.8	37.7	36.8	59	43	40	52	33	35	44	31	32	46	42	31	55	37	39	55	38	37	55	36	40
Change After-Before	12 3	2 9	2.8	12	0	8	11	5	1	13	6	· 4	25	3	7	10	−4	6	10	8	1	6	3	−3
Significance / x^2 (1df)																								
Adelaide *v.* Hobart	V		32 1	H		7.0	V		15.0	V		16.6	V		28 0	.		2 7	S		6 2	H		6.7
Melbourne *v.* Hobart	.		0 0	H		8 0	.		2 4	S		5.3			2 1	H		7 6	.		2.8	.		2.3
Adelaide *v.* Melbourne	V		36.0	V		24 0	H		7.8	.		2.7	V		40.9	V		19.0	.		1 6	.		2.9
FLATNESS																								
No twist—																								
Base Level Before (%)	75.0	73 8	76.0	78	74	80	76	74	76	68	71	69	70	76	75	76	76	75	76	70	73	77	74	80
Change After-Before	6.3	1.8	−0 1	6	3	−2	5	1	2	10	1	1	13	−1	−6	4	−1	−1	5	9	5	5	−1	−1
Significance/x^2 (1df)—																								
Adelaide *v.* Hobart	V		17.8	H		10.1	.		1.5	S		5 9	V		30.2	.		1.8	.		0.1	.		3.4
Melbourne *v.* Hobart			1 4	S		4 3	.		0.1	.		0 0	.		2 0	.		0.0	.		1.9	.		0 0
Adelaide *v.* Melbourne	V		11 6	.		2.5	.		3.4	S		5.7	V		19.0	.		2.9	.		1.3	S		4.4
No. observation (average)																								
Before publicity	3 388	3 869	2 795	1 174	1 777	1 223	1 482	1 607	1 033	706	474	534	744	638	631	770	1 084	730	812	1 014	758	1 062	1 134	675
After publicity	3 411	3 106	2 092	1 305	1 361	893	1 456	1 346	750	632	395	448	825	786	414	777	721	514	760	803	585	1 049	795	579

Key: V significant at 0.1% level ($p \leqslant 0.001$)
H significant at 1% level ($0.01 \geqslant p > 0.001$)
S significant at 5% level ($0.05 \geqslant p > 0.01$)
. not significant ($p > 0.05$)

Table II: Results for Experiment 1 controlled for time of day

Experiment 1	*All occupants*			*6.00–8.59 a.m.*			*9.00–11.59 a.m.*			*12.00–2.59 p.m.*			*3.00–5.59 p.m.*			*6.00–8.59 p.m.*			*9.00–11.59 p.m.*		
City	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*	*Adel.*	*Melb.*	*Hob.*
TIGHTNESS																					
Tight or In-between—																					
Base Level Before (%)	61.0	57.1	67.2	73	59	71	66	55	63	51	54	72	58	57	68	64	58	70	53	63	61
Change After-Before	8.7	3.0	2.8	15	−2	−5	13	0	8	9	3	2	4	8	−3	2	4	−2	13	1	18
Significance / χ^2 *(1df)*—																					
Adelaide *v.* Hobart	H		10.4	V		20.5	.		3.6	.		2.4	S		4.3	.		0.7	.		1.7
Melbourne *v.* Hobart	.		0.2	.		0.6	S		4.9	.		0.2	H		7.4	.		2.2	H		9.1
Adelaide *v.* Melbourne	V		14.0	V		24.2	V		15.4	.		2.1	.		1.1			0.3	S		4.0
BUCKLE POSITION																					
Off-hip—																					
Base Level Before (%)	52.8	37.7	36.8	63	34	37	61	37	34	41	36	43	49	38	34	59	45	39	39	37	39
Change After-Before	12.3	2.9	2.8	14	7	−12	9	6	7	16	6	2	11	9	2	5	−7	2	27	−3	5
Significance / χ^2 *(1df)*—																					
Adelaide *v.* Hobart	V		32.1	V		24.1	S		4.9	H		8.6	H		7.2	.		1.8	H		9.8
Melbourne *v.* Hobart	.		0.0	V		10.9	.		0.1	.		0.7	.		2.5	.		3.4	.		1.9
Adelaide *v.* Melbourne	V		36.0	H		7.2	S		4.0	S		6.1	.		1.7	H		8.7	V		24.5
FLATNESS																					
No twist—																					
Base Level Before (%)	75.0	73.8	76.0	77	81	73	72	73	73	73	70	75	74	72	77	82	80	80	71	68	80
Change After-Before	6.3	1.8	−0.1	7	−2	4	10	3	4	5	2	0	5	5	−4	0	2	−5	13	−1	2
Significance / χ^2 *(1df)*—																					
Adelaide *v.* Hobart	V		17.8	.		1.0	S		4.8	.		1.9	H		9.4	.		1.3	.		3.5
Melbourne *v.* Hobart	.		1.4	.		1.2	.		0.1	.		0.3	S		6.0	.		2.3	.		0.2
Adelaide *v.* Melbourne	V		11.6	S		6.4	S		5.3	.		1.0	.		0.3	.		0.1	H		8.0
No. observations (average)—																					
Before publicity	3 388	3 869	2 795	522	553	239	599	567	805	600	789	426	1 012	795	690	485	724	433	170	440	202
After publicity	3 411	3 106	2 092	480	398	172	725	452	644	532	635	386	1 025	560	494	420	702	221	228	356	175

Key: V significant at 0.1% level ($p \leqslant 0.001$)
H significant at 1% level ($0.01 \geqslant p > 0.001$)
S significant at 5% level ($0.05 \geqslant p > 0.01$)
. not significant ($p > 0.05$)

APPENDIX C

Detailed results—Experiment 2

Need for controlled analysis

Only time of day, occupant age and day of week were investigated as potential 'controlling' descriptors for Experiment 2. In Hobart the apparent effect of the publicity on tightness was found to vary significantly with time of day (< 0.05) and on buckle position with time of day, age and day of week (all < 0.05). The distribution of observations by time of day and day of week also varied significantly (< 0.01) between the before and after stages in Hobart. The control cities were not investigated. From these analyses it did not appear imperative to control for occupant age, but again a conservative stance was adopted and a decision made to control for time of day, day of week and occupant age for all three measures of correctness.

The results from the controlled analyses are presented in full in Tables I and II and are summarised below, reporting particular results which appear to be strong exceptions to the overall publicity effects in Hobart. The use of two control cities gives better information on the changes expected to occur without any publicity treatment (though the treatment and control subjects are different, at very least geographically). An apparent exceptional publicity effect was only thought to be truly so if the Hobart changes relative to both control cities were exceptional.

Summary of controlled analyses

The following results are exceptions to the effects of a high level of publicity on:

- 'tight' (relative change 18.9 per cent of Melbourne, 15.6 per cent of Adelaide, for all occupants)
  - 6.00 to 8.59 p.m. (+ 24 per cent cf. Melbourne, +24 per cent cf. Adelaide)
  - 9.00 to 11.59 p.m. (+ 10 per cent cf. Melbourne, + 3 per cent cf. Adelaide)
- buckle 'off-hip' (relative change 13.8 per cent cf. Melbourne, 10.7 per cent cf. Adelaide, for all occupants)
  - Sunday (+ 9 per cent cf. Melbourne, + 2 per cent cf. Adelaide)
  - 6.00 to 8.59 a.m. (+ 23 per cent cf. Melbourne, + 35 per cent cf. Adelaide)
  - 9.00 to 11.59 p.m. (– 1 per cent cf. Melbourne, + 5 per cent cf. Adelaide)
- 'no twist' (relative change 7.1 per cent cf. Melbourne, 3.8 per cent cf. Adelaide, for all occupants)
  - Saturday (+ 3 per cent cf. Melbourne, – 5 per cent cf. Adelaide)
  - 6.00 to 8.59 a.m. (+ 17 per cent cf. Melbourne, + 15 per cent cf. Adelaide)
  - 9.00 to 11.59 p.m. (+ 3 per cent cf. Melbourne, zero cf. Adelaide)

Table I: Results for Experiment 2 controlled for age and day of week

Experiment 2	*All occupants*			*Age 8–29*			*Age 30–49*			*Age 50+*		
City	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*
TIGHTNESS												
Tight or In-between—												
Base Level Before (%) .	48.2	57 3	56.2	51	61	61	47	55	56	44	54	45
Change After-Before .	15.4	−3.5	−0.2	16	−5	0	16	−2	0	14	−6	−2
Significance/χ^2 (1df)—												
Hobart v. Melbourne .	V		112.6	V		60.4	V		38 2	V		18.2
Hobart v. Adelaide .	V		70.8	V		34 3	V		24.2	V		13 4
Melbourne v. Adelaide .	.		3.7	.		3.4	.		0.8	.		0.6
BUCKLE POSITION												
Off-hip—												
Base Level Before (%) .	37 1	49 3	54.6	44	53	62	33	48	50	27	45	47
Change After-Before .	18.2	4 4	7.5	21	3	6	16	7	9	16	− 3	8
Significance /χ^2 (1df) .												
Hobart v. Melbourne .	V		60.1	V		39.1	V		12.6	V		20 0
Hobart v. Adelaide .	V		34.6	V		25 0	H		7 5	S		5.2
Melbourne v. Adelaide .	.		3.7	.		1.5	.		0.7	S		6.3
FLATNESS												
No twist—												
Base Level Before (%) .	74.0	75.8	79.3	76	77	82	73	75	79	71	74	73
Change After-Before .	8.2	1.1	4.4	8	2	5	9	0	3	6	1	5
Significance /χ^2 (1df)—												
Hobart v. Melbourne .	V		24.3	H		9.2	V		14.4	.		1.7
Hobart v. Adelaide .	S		4.4	.		1.3	S		3 9	.		0.1
Melbourne v. Adelaide .	H		7.3	.		3.4	.		2.5	.		1.2
No. observation (average)—												
Before publicity . .	2 584	3 696	3 092	1 145	1 399	1 342	961	1 726	1 173	478	567	567
After publicity . .	2 741	3 773	3 052	1 262	1 518	1 346	968	1 781	1 143	509	471	552

Experiment 2	*Thursday*			*Friday*			*Saturday*			*Sunday*		
City	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*
TIGHTNESS												
Tight or In-between—												
Base Level Before (%) .	48	58	51	45	58	51	46	54	54	53	59	66
Change After-Before .	14	−8	1	17	−7	2	17	8	5	14	−6	−4
Significance/χ^2 (1df)—												
Hobart v. Melbourne .	V		35.4	V		40 1	H		7.0	V		37.1
Hobart v. Adelaide .	V		12.1	V		15.0	H		10.3	V		27.9
Melbourne v. Adelaide .	S		6.5	H		6 7	.		0.7	.		0,5
BUCKLE POSITION												
Off-hip—												
Base Level Before (%) .	34	48	59	36	48	60	33	52	54	44	50	49
Change After-Before .	26	8	4	21	− 2	1	17	7	12	13	4	11
Significance /χ^2 (1df) .												
Hobart v. Melbourne .	V		23.6	V		35.4	H		9.9	H		7.1
Hobart v. Adelaide .	V		31.6	V		25.5	S		3.9	.		0 6
Melbourne v. Adelaide .	.		1.2	.		0.8	.		2.4	S		4.7
FLATNESS												
No twist—												
Base Level Before (%) .	68	79	77	69	74	82	79	82	77	79	70	81
Change After-Before .	11	0	2	9	− 3	− 1	3	0	8	9	5	10
Significance /χ^2 (1df)—												
Hobart v. Melbourne .	H		10.0	V		12.8	.		1.2	S		5.1
Hobart v. Adelaide .	H		7.5	H		9 0	.		3.3	.		1.8
Melbourne v. Adelaide .	.		0.1	.		0.2	H		8.6	V		13.5
No. observation (average)—												
Before publicity . .	619	919	707	613	783	760	611	896	615	741	1 098	1 010
After publicity . .	551	962	818	568	803	728	838	953	689	783	1 055	818

Key:
- V significant at 0.1% level ($p \leqslant 0.001$)
- H significant at 1% level ($0\ 01 \geqslant p > 0.001$)
- S significant at 5% level ($0.05 \geqslant p > 0.01$)
- . not significant ($p > 0.05$)

Table II: Results for Experiment 2 controlled for time of day

Experiment 2	*All occupants*			*6.00–8.59 a.m.*			*9.00–11.59 a.m.*			*12.00–2.59 p.m.*			*3.00–5.59 p.m.*			*6.00–8.59 p.m.*			*9.00–11.59 p.m.*		
City	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*	*Hob.*	*Melb.*	*Adel.*
TIGHTNESS																					
Tight or In-between—																					
Base Level Before (%)	48.2	57.3	56 2	44	57	56	51	55	54	56	61	53	39	64	56	51	53	65	53	54	55
Change After-Before	15.4	−3 5	−0.2	15	−7	−2	14	−1	5	13	−8	−1	11	−4	−4	22	−2	−2	11	1	8
Significance χ^2 *(1df)—*																					
Hobart *v.* Melbourne	V		112.6	V		12.6	V		17 2	V		22.0	V		14.5	V		35.6	.		3.1
Hobart *v.* Adelaide	V		70 8	H		6.8	S		5.3	H		9.7	V		13 8	V		28.6	.		0.5
Melbourne *v.* Adelaide	.		3.7	.		1.1	.		2.2			3.1	.		0 1	.		0 1	.		1.5
BUCKLE POSITION																					
Off-hip—																					
Base Level Before (%)	37.1	49.3	54.6	22	49	48	31	51	43	36	53	40	41	53	54	39	52	64	52	37	91
Change After-Before	18.2	4.4	7.5	32	9	−3	17	4	6	19	0	14	17	5	7	21	1	13	10	11	5
Significance χ^2 *(1df)—*																					
Hobart *v.* Melbourne	V		60.1	V		16 1	V		12.5	V		19.6	H		9.1	V		26.4	.		0.4
Hobart *v.* Adelaide	V		34.6	V		30.7	H		7 4			1 8	S		5.9	S		4 6	.		3.3
Melbourne *v.* Adelaide	.		3.7	S		6.5	.		0.3	V		11 4	.		0 6	V		12.1	S		5 8
FLATNESS																					
No twist—																					
Base Level Before (%)	74 0	75.8	79.3	70	77	74	75	79	79	77	74	78	68	75	75	77	74	82	79	77	94
Change After-Before	8.2	1.1	4.4	14	−3	−1	6	−2	−7	6	6	5	12	0	8	7	4	11	5	2	5
Significance χ^2 *(1df)—*																					
Hobart *v.* Melbourne	V		24.3	V		10.9	S		6.3	.		0.2	V		11.4	.		1.2	.		0 4
Hobart *v.* Adelaide	S		4.4	H		8.0	V		13.5	.		0.1	.		0.5	H		6.9	S		6.4
Melbourne *v.* Adelaide	H		7.3	.		0.2	.		1.8	.		0.1	H		9 3	V		13.1	H		9.4
No. observations (average)—																					
Before publicity	2 584	3 696	3 092	152	565	386	657	643	491	407	598	537	663	683	865	505	673	466	200	533	347
After publicity	2 741	3 773	3 052	203	576	384	764	671	456	443	597	641	508	652	789	565	730	412	259	548	371

Key:
- V significant at 0.1% level ($p \leqslant 0.001$)
- H significant at 1% level ($0.01 \geqslant p > 0.001$)
- S significant at 5% level ($0.05 \geqslant p > 0.01$)
- . not significant ($p > 0.05$)

APPENDIX D

Detailed results—Experiment 3

Controlled analysis

In Experiment 3, the five descriptors of each occupant observed were augmented by another two:

- wearing overcoat (yes/no)
- weather

These were added since the measurements were taken in winter and the range of climates across cities was considerable.

Where controlled analysis was found necessary for each of the four experimental cities, this is indicated on the appropriate table (Tables I-IV) for that city and the results of controlled analyses given. Empty sections of the tables indicated that it was not necessary to control for the particular descriptor when analysing the particular measure of correctness.

Summary of controlled analyses

The results from the controlled analyses are discussed for each treatment city in turn, reporting particular results which appear to be strong exceptions to the overall publicity effects found for all occupants.

Newcastle (HIGH-SHORT): The following results are exceptions to the effects of high-short publicity on:

- 'tight' or 'in-between' (relative change cf. CONTROL, 3.0 per cent for all occupants)
  - 6.00 to 8.59 p.m. (+11 per cent relative change)
  - Thursday (+11 per cent)
  - Saturday (+16 per cent)
  - sunny weather (+8 per cent)
- buckle 'off-hip' (relative change 4.4 per cent for all occupants)
  - Thursday (−4 per cent)
  - Saturday (+15 per cent)
  - driver with passenger (+7 per cent)
  - sunny weather (+ 12 per cent)
  - overcast weather (+ 10 per cent)
- 'no twist' (relative change 0.1 per cent for all occupants)
  - 6.00 to 8.59 a.m. (− 9 per cent)
  - 12.00 to 2.59 p.m. (+ 10 per cent)
  - 9.00 to 10.59 p.m. (− 15 per cent)
  - Thursday (− 7 per cent)
  - Saturday (+ 10 per cent)
  - driver with passenger (+ 6 per cent)

Wollongong (HIGH-LONG): The following results are exception to the effects of high-long publicity on:

- 'tight' or 'in-between' (relative change cf. CONTROL, 2.2 per cent for all occupants)
  - 12.00 to 2.59 p.m. (− 11 per cent)
  - 6.00 to 8.59 p.m. (+ 16 per cent)
- buckle 'off-hip' (relative change 0.1 per cent for all occupants)
  - Friday (+ 9 per cent)
  - sunny weather (+ 6 per cent)
- 'no twist' (relative change 4.6 per cent for all occupants)
  - Friday (+ 16 per cent)
  - sunny weather (+ 9 per cent)

Brisbane (LOW-SHORT): The following results are exceptions to the effects of low-short publicity on:

- 'tight' or 'in-between' (relative change cf. CONTROL, 7.3 per cent for all occupants)
  - Thursday (+ 12 per cent)
  - Saturday (+ 15 per cent)
  - sunny weather (+ 18 per cent)
- buckle 'off-hip' (relative change − 0.9 per cent for all occupants)
  - 6.00 to 8.59 a.m. (+ 15 per cent)
  - 6.00 to 8.59 p.m. (− 13 per cent)
- 'no twist' (relative change 2.7 per cent for all occupants)
  - 12.00 to 2.59 p.m. (+ 12 per cent)
  - overcast weather (+ 8 per cent)

Perth (LOW-LONG): The following results are exceptions to the effects of low-long publicity on:

- 'tight' or 'in-between' (relative change − 0.4 per cent for all occupants)
  - 6.00 to 8.59 a.m. (+ 14 per cent)
  - 12.00 to 2.59 p.m. (− 17 per cent)
  - Saturday (+ 8 per cent)
- buckle 'off-hip' (relative change 0.6 per cent for all occupants)
  - 6.00 to 8.59 a.m. (+ 13 per cent)
  - Thursday (+ 8 per cent)
- 'no twist' (relative change 4.6 per cent for all occupants)
  - 6.00 to 8.59 a.m. (+ 9 per cent)
  - 9.00 to 11.59 a.m. (+ 12 per cent)
  - 12.00 to 2.59 p.m. (+ 13 per cent)
  - Saturday (+ 8 per cent)
  - wearing overcoat (+ 16 per cent)

Table Ia: Results for Experiment 3 from Newcastle and Sydney

Newcastle (high-short)	*All occupants*		*Male*		*Female*		*6.00–8.59 a.m.*		*9.00–11.59 a.m.*		*12.00–2.59 p.m.*		*3.00–5.59 p.m.*		*6.00–8.59 p.m.*		*9.00–10.59 p.m.*	
City	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*
TIGHTNESS																		
Tight on In-between																		
Base Level Before (%)	67.9	63.1	70	64	64	61	70	50	65	49	65	56	64	76	75	78	79	77
Change After-Before	7.6	4.6	6	4	9	7	8	8	10	16	5	1	10	3	4	−7	2	0
Significance	S		.		.		.		.		.		.		H		.	
x^2 *(1df)*	4.7		3.5		0.7		1.0		2.1		0.6		2.8		8.6		0.1	
BUCKLE POSITION																		
Off-hip																		
Base Level Before (%)	41.1	48.5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Change After-Before	12.9	8.5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Significance	S																	
x^2 *(1df)*	6.3		..		..		..		..		..		..		.		..	
FLATNESS																		
No twist																		
Base Level Before (%)	80.2	82.0	..	..	..	..	86	76	78	78	74	79	78	85	82	89	89	88
Change After-Before	0.8	0.7	..	..	..	..	−6	3	4	3	7	−3	2	1	1	−2	−14	1
Significance	.		.		.		H		.		S		.		.		S	
x^2 *(1df)*	0.0		..		..		7 0		0.2		5.8		0.1		1.3		5 6	
No observations (average)																		
Before publicity	3 039	2 501	1 898	1 659	1 145	855	523	431	713	538	344	385	766	502	555	457	140	189
After publicity	3 462	2 440	2 297	1 610	1 170	835	487	431	614	486	403	356	1 078	553	711	472	167	144

Key: V significant at 0.1% level ($p \leqslant 0.001$)
H significant at 1% level ($0.01 \geqslant p > 0.001$)
S significant at 5% level ($0.05 \geqslant p > 0.01$)
. not significant ($p > 0.05$)

Table Ib: Results for Experiment 3 from Newcastle and Sydney

Newcastle (high-short)		*Thursday*		*Friday*		*Saturday*		*Sunday*		*Driver alone*		*Driver with passenger*		*Passenger*		*Sunny*		*Overcast*		*Rainy*	
	City	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*	*New.*	*Syd.*
TIGHTNESS																					
Tight or In-between																					
Base Level Before (%) . .		64	67	72	51	63	68	70	64	..	..	..	.	..	..	72	69	60	53	73	69
Change After-Before . .		13	2	0	16	14	−2	6	5	..	..	..	..	..	..	5	−3	11	17	10	8
Significance— x^2 *(1df)*		V 11.8		V 17.1		V 19.1		. 0.3								H 10.0		. 3.3		. 0.6	
BUCKLE POSITION																					
Off-hip—																					
Base Level Before (%) . .		50	49	42	49	35	41	39	52	49	49	40	52	36	44	32	50	46	48	44	45
Change After-Before . .		2	6	10	7	25	10	15	12	9	9	15	8	14	8	18	6	17	7	−1	21
Significance— x^2 *(1df)*		. 0 9		. 0.9		V 13.3		. 1.6		. 0.3		S 4.3		. 2.8		V 22.5		H 9.3		H 10 3	
FLATNESS																					
No twist—																					
Base Level Before (%) . .		80	80	79	80	76	85	82	83	78	78	81	83	81	86	81	83	81	80	79	85
Change After-Before . .		−5	2	1	3	8	−2	3	−1	−3	1	5	−1	2	3	−2	−2	4	6	8	5
Significance— x^2 *(1df)*		S 4.4		. 0 5		H 8.4		. 1.8		. 1 7		S 5.1		. 0 3		. 0.1		. 0.6		. 0.5	
No. observations (average)—																					
Before publicity		687	734	512	509	500	528	1 336	727	916	989	1 113	793	1 009	713	906	1 065	1 145	916	983	517
After publicity		869	583	905	638	704	573	977	646	1 349	1 023	1 108	733	1 003	682	2 252	1 701	1 100	592	104	147

Key: V significant at 0.1% level ($p \leqslant 0.001$)
H significant at 1% level ($0.01 \geqslant p > 0.001$)
S significant at 5% level ($0.05 \geqslant p > 0.01$)
. not significant ($p > 0.05$)

Table IIa: Results for Experiment 3 from Wollongong and Sydney

Wollongong (high-long)	*City*	*All occupants* Woll.	Syd.	*6.00–8.59 a.m.* Woll.	Syd.	*9.00–11.59 a.m.* Woll.	Syd.	*12.00–2.59 p.m.* Woll.	Syd.	*3.00–5.59 p.m.* Woll.	Syd.	*6.00–8.59 p.m.* Woll.	Syd.	*9.00–10.59 p.m.* Woll.	Syd.
TIGHTNESS															
Tight or In-between															
Base Level Before (%)	. .	62.2	60.1	67	60	68	49	77	45	62	73	49	67	46	65
Change After-Before	. .	9.9	7.7	7	− 2	5	15	2	13	8	7	20	4	16	11
Significance x^2 *(1df)*	. . . . .	. 2.1		S 5.3		S 4.9		S 4.5		. 0.8		H 10.7		. 0.4	
BUCKLE POSITION															
Off-hip—															
Base Level Before (%)	. .	48.3	45.6												
Change After-Before	. .	11.5	11.4												
Significance x^2 *(1df)*	. . . . .	. 0.1													
FLATNESS															
No twist—															
Base Level Before (%)	. .	77.7	82.5	75	80	74	81	82	84	75	82	83	84	80	87
Change After-Before	. .	4.8	0.2	5	− 1	7	0	− 2	− 8	12	4	2	3	12	2
Significance x^2 *(1df)*	. . . . .	H 7.7		. 2.0		. 3.3		. 2.3		S 4.2		. 0.2		. 2.6	
No. observations (average)—															
Before publicity	. . . .	2 505	2885	437	538	456	544	337	449	621	650	508	519	147	182
After publicity	. . . .	2 213	2440	551	431	454	486	374	356	322	553	395	472	113	144

Key: V significant at 0.1% level ($p \leq 0.001$)
H significant at 1% level ($0.01 \geq p > 0.001$)
S significant at 5% level ($0.05 \geq p > 0.01$)
. not significant ($p > 0.05$)

Table IIb: Results for Experiment 3 from Wollongong and Sydney

Wollongong (high-long)	*City*	*Thursday* Woll.	*Thursday* Syd.	*Friday* Woll.	*Friday* Syd.	*Saturday* Woll.	*Saturday* Syd.	*Sunday* Woll.	*Sunday* Syd.	*Overcoat* Woll.	*Overcoat* Syd.	*No overcoat* Woll.	*No overcoat* Syd.	*Sunny* Woll.	*Sunny* Syd.	*Overcast* Woll.	*Overcast* Syd.	*Rainy* Woll.	*Rainy* Syd.
TIGHTNESS																			
Tight or In-between—																			
Base Level Before (%)		..	..	..	..	..	..	..	..	..	..	..	..	60	55	61	68	72	59
Change After-Before		..	..	..	..	..	..	..	.	..	..	..	..	13	12	9	2	− 22	17
Significance x^2 *(1df)*		..		..		..		..		..		..		. 3.5		S 4.2		H 10 5	
BUCKLE POSITION																			
Off-hip																			
Base Level Before (%)		45	46	47	48	46	43	54	45	41	41	49	44	43	45	54	47	46	46
Change After-Before		7	9	17	8	15	9	13	19	5	8	13	14	18	12	4	9	− 31	20
Significance x^2 *(1df)*		. 0 5		S 5.8		2 3		3.4		. 0 1		. 0.7		S 6.4		. 1.3		V 15.0	
FLATNESS																			
No twist-																			
Base Level Before (%)		77	80	69	84	81	82	82	84	..	..	..		75	82	83	82	72	88
Change After-Before		1	3	15	− 1	5	1	5	− 1	..	.	..	..	8	− 1	1	4	21	2
Significance x^2 *(1df)*		. 0 5		V 18.7		. 2 0		. 3.8		. ..		. ..		V 14 5		. 1.0		H 7.2	
No observations (average)—																			
Before publicity		698	743	454	694	618	639	731	800	139	225	2 341	2 538	922	1 483	1 126	1 067	454	331
After publicity		692	582	608	637	592	573	316	646	236	365	1 955	2 044	1 804	1 701	390	592	14	147

Key: V significant at 0.1% level ($p \leq 0.001$)
H significant at 1% level ($0.01 \geq p > 0.001$)
S significant at 5% level ($0.05 \geq p > 0.01$)
. not significant ($p > 0.05$)

Table IIIa: Results for Experiment 3 from Brisbane and Sydney

Brisbane (low-short)	*All occupants*		*6.00–8.59 a.m.*		*9.00–11.59 a.m.*		*12.00–2.59 p.m.*		*3.00–5.59 p.m.*		*6.00–8.59 p.m.*		*9.00–10.59 p.m.*	
City	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*
TIGHTNESS														
Tight or In-between—														
Base Level Before (%)	52.7	63.1	..	..	..	..	..	..	..	..	..	..	..	..
Change After-Before	11.9	4.6	..	..	..	..	..	..	..	..	..	..	..	..
Significance	V		..		..		..		..		..		..	
x^2 *(1df)*	12.7		..		..		..		..		..		..	
BUCKLE POSITION														
Off-hip—														
Base Level Before (%)	51.7	48.5	49	48	53	45	46	48	47	56	61	47	55	43
Change After-Before	7.6	8.5	16	1	13	8	7	2	10	14	−3	10	10	17
Significance	.		H		.		.		.		H		.	
x^2 *(1df)*	0.2		8.6		2.1		1.0		2.1		8.8		1.1	
FLATNESS														
No twist—														
Base Level Before (%)	79.0	82.0	77	76	79	78	78	79	77	85	82	89	82	88
Change After-Before	3.4	0.7	3	3	4	3	9	−3	8	1	−3	−2	−4	1
Significance	.		.		.		V		.		.		.	
x^2 *(1df)*	2.9		0.0		0.1		11.6		4.2		0.0		0.7	
No. observations (average)—														
Before publicity	2 603	2 501	358	428	439	537	479	384	538	500	607	455	181	189
After publicity	2 819	2 440	319	430	475	485	484	356	686	552	634	471	215	145

Key: V significant at 0.1% level ($p \leqslant 0.001$)
H significant at 1% level ($0.01 \geqslant p > 0.001$)
S significant at 5% level ($0.05 \geqslant p > 0.01$)
. not significant ($p > 0.05$)

Table IIIb: Results for Experiment 3 from Brisbane and Sydney

Brisbane (low-short)		*Thursday*		*Friday*		*Saturday*		*Sunday*		*Overcoat*		*No overcoat*		*Sunny*		*Overcast*	
	City	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*	*Bris.*	*Syd.*
TIGHTNESS																	
Tight or In-between—																	
Base Level Before (%)		47	67	54	51	52	38	59	64	56	65	53	63	49	69	54	53
Change After-Before		14	2	9	16	13	−2	10	5	10	5	12	5	15	−3	16	17
Significance— x^2 *(1df)*		H 9.7		S 3.9		V 14.2		2.0		0.6		V 10.3		V 49.9		2.7	
BUCKLE POSITION																	
Off-hip—																	
Base Level Before (%)		51	49	60	49	51	42	46	52	..	..	..	..	51	50	55	48
Change After-Before		11	6	0	7	12	10	8	12	..	..	..	..	8	6	9	7
Significance— x^2 *(1df)*		1.5		2.8		0.6		1.8		..		..		0.7		1.1	
FLATNESS																	
No twist—																	
Base Level Before (%)		..	..	..	..	..	..	..	..	..	..	..	..	81	83	78	80
Change After-Before		..	..	..	..	..	..	..	..	..	..	..	..	1	−2	14	6
Significance— x^2 *(1df)*		..		..		..		..		..		..		1.8		H 7.6	
No. observations (average)—																	
Before publicity		697	734	620	510	678	529	608	729	158	299	2 439	2 159	1 416	1 065	810	916
After publicity		621	582	722	640	773	574	698	646	327	367	2 479	2 053	2 622	1 701	191	592

Key: V significant at 0.1% level ($p \leq 0.001$)
H significant at 1% level ($0.01 \geq p > 0.001$)
S significant at 5% level ($0.05 \geq p > 0.01$)
. not significant ($p > 0.05$)

Table IVa: Results for Experiment 3 from Perth and Sydney

Perth (low-long)		*All occupants*		*Male*		*Female*		*6.00–8.59 a.m.*		*9.00–11.59 a.m.*		*12.00–2.59 p.m.*		*3.00–5.59 p.m.*		*6.00–8.59 p.m.*		*9.00–10.59 p.m.*	
	City	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*	*Perth*	*Syd.*
TIGHTNESS																			
Tight or In-between																			
Base Level Before (%)		58.7	60.1	..	..	..	..	61	60	51	49	56	45	59	73	62	67	63	65
Change After–Before		7.3	7.7	..	..	..	..	12	−2	12	15	−4	13	10	7	8	4	7	11
Significance				..		..		V		.		V		.		.		.	
x^2 *(1df)*		0.4		..		..		11.6		1.2		13.5		0.8		0 9		0.7	
BUCKLE POSITION																			
Off-hip—																			
Base Level Before (%)		41.0	45.6	41	46	41	44	42	41	37	42	30	40	40	52	47	50	59	48
Change After–Before		12.0	11.4	14	12	8	10	21	8	12	11	4	9	13	19	14	7	3	12
Significance		.		.		.		H		.		.		S		.		.	
x^2 *(1df)*		1.3		2.0		0.9		9.4		0.6		1.7		4.4		3.3		1.7	
FLATNESS																			
No twist—																			
Base Level Before (%)		72.5	82.5	..	..	..	..	74	80	68	81	68	84	72	82	78	84	71	87
Change After–Before		4.8	0.2	..	..	..	..	8	−1	12	0	5	−8	5	4	0	3	4	2
Significance		H		..		..		S		H		V		.		.		.	
x^2 *(1df)*		8.2		..		..		5.5		9.7		11.4		0.3		1.1		0.5	
No. observations (average)—																			
Before publicity		3 730	2 885	2 491	1 938	1 237	946	483	537	514	545	562	449	1 027	649	958	517	182	183
After publicity		4 101	2 440	2 642	1 608	1 458	829	556	430	721	485	642	356	1 071	552	853	472	256	145

Key:
- V significant at 0.1% level ($p \leqslant 0.001$)
- H significant at 1% level ($0.01 \geqslant p > 0.001$)
- S significant at 5% level ($0.05 \geqslant p > 0.01$)
- . not significant ($p > 0.05$)

Table IVb: Results for Experiment 3 from Perth and Sydney

Perth (low-long)	Thursday		Friday		Saturday		Sunday		Overcoat		No overcoat		Sunny		Overcast	
City	Perth	Syd.	Perth	Syd.	Perth	Syd.	Perth	Syd.	Perth	Syd.	Perth	Syd.	Perth	Syd.	Perth	Syd.
TIGHTNESS																
Tight or In-between—																
Base level before (%)	54	52	62	61	60	66	60	62	..	..	.	..	57	55	59	68
Change After-Before	13	17	4	7	8	0	4	7	..	..	..	.	9	12	3	2
Significance	.		.		S		.		.				.		.	
χ^2 (*1df*)	1.7		0.8		5.0		1.0		.		..		2.2		0 9	
BUCKLE POSITION																
Off-hip—																
Base level before (%)	41	46	43	48	42	43	39	45	..	..	..	..	42	45	39	47
Change After-Before	17	9	6	8	11	9	15	19	..	..	..	..	11	12	14	9
Significance	S		.		.				..		..		.		.	
χ^2 (*1df*)	4.6		0.4		0.5		2.8		..		..		1.6		2.8	
FLATNESS																
No twist—																
Base level before (%)	72	80	75	84	69	82	73	84	61	80	73	83	..	.	..	.
Change After-Before	2	3	3	−1	9	1	6	−1	16	0	5	0	..	..	..	..
Significance			.		S		S		S		S		..		..	
χ^2 (*1df*)	0 2		1.5		5 9		6.1		5.5		5 8		..		.	
No. observations (average)—																
Before publicity	1 135	743	919	696	812	640	860	802	114	226	3 584	2 618	2 224	1 487	886	1 067
After publicity	1 081	583	999	638	1 089	573	930	646	245	366	3 839	2 048	3 907	1 700	190	594

Key:
- V significant at 0.1% level ($p \leq 0.001$)
- H significant at 1% level ($0.01 \geq p > 0.001$)
- S significant at 5% level ($0.05 \geq p > 0.01$)
- . not significant ($p > 0.05$)