



**27th Australasian Transport Research Forum, Adelaide, 29 September – 1 October 2004**

**Paper title:** Evaluating voluntary travel behaviour interventions

**Author(s) name(s):** Peter Stopher, Rahaf Alsnih, Philip Bullock, and Elizabeth Ampt

**Organisation(s):** Institute of Transport Studies, The University of Sydney (first three) and Steer Davies Gleave (fourth author)

**Contact details:**

*Postal address:* Institute of Transport Studies, C37, The University of Sydney, NSW 2006

*Telephone:* (02) 9351 0010

*Facsimile:* (02) 9351 0088

*email:* [peters@its.usyd.edu.au](mailto:peters@its.usyd.edu.au)

**Abstract (200 words):**

Considerable interest in the policy of voluntary travel behaviour change interventions, known as by the generic name of TravelSmart®, has emerged. Measuring its effectiveness and determining its cost-benefit ratios is a major issue. Several difficulties arise in this process. First, it requires both a before and an after survey, sufficiently far apart to detect stable change in household travel behaviour. Second, it requires estimates of numbers of trips and activities, distance travelled by mode, time spent travelling by mode, and the modes of travel used. These are poorly reported in household travel surveys, introducing serious potentials for error in evaluation. Third, are issues relating to sample sizes to detect changes of the order of 5 to 10 percent in various travel behaviours, with acceptable accuracy. After discussing these issues in some detail, we describe a potential survey process, using GPS devices, that can overcome a number of the problems. We describe the information that can be obtained through the GPS and its associated prompted recall survey, demonstrating some of the benefits associated with this procedure. We conclude that evaluation is a significant issue that requires substantial funding to be done effectively, but that the GPS survey offers a high level of reliability in the information obtained.

### Introduction

Over the past seven years, there has emerged a considerable interest in Australia in the policy of voluntary travel behaviour change interventions. These are best known under the generic name of TravelSmart®. Although applications of these policies are starting to be seen in other countries around the world, Australia is probably the pioneer of the application of this strategy, with efforts underway in all of the capital regions of the country. One of the major issues with the strategy is, however, measuring its effectiveness, and also determining the cost-benefit ratio associated with it as a policy.

There are a number of difficulties that arise in this process. First, it requires a before and an after survey, spaced sufficiently far apart to detect stable change in household travel behaviour. Second, ideally it requires estimates of numbers of trips, numbers of activities, total distance travelled by mode, total time spent travelling by mode, and the modes of travel used, to estimate the changes in these values from before the intervention to after. Almost all of these are notoriously poorly reported in household travel surveys, thereby introducing serious potentials for error in the evaluation measurement. Third, there are serious issues relating to the sample sizes needed for such before and after surveys, to detect changes of the order of 5 to 10 percent in various travel behaviours, with an acceptable level of accuracy. Fourth, in many cases there are many beneficial changes other than those related directly to transport (e.g., social, health and community benefits). Methods for evaluating these changes are not well developed.

This paper discusses the first three issues in some detail, providing evidence of some of the problems inherent in the evaluation measurements. The fourth issue is discussed in Ampt, 2001. We then describe a potential survey process, using GPS devices, that can overcome a number of the potential problems in evaluation discussed in the paper. We show the information that can be gained through the GPS survey and its associated prompted recall, and demonstrate some of the benefits associated with this procedure. We conclude that evaluation is a significant issue that requires substantial funding to be able to be done effectively, but that methods such as the GPS survey will provide a high level of reliability in the information obtained.

### Overview of Voluntary Travel Behaviour Change

Fundamentally, Voluntary Travel Behaviour Change is based on the notion of seeking to generate *voluntary* change in travel behaviour. Some may think that any changes in travel behaviour are voluntary. However, the specific notion of “voluntary” in this application is that nothing is changed in the transport system, no pricing or taxing incentives or disincentives are provided, and it is rather a matter of providing better information to people about their transport (and non-transport) options. The idea is that, given this improved level of information, people will decide to make changes in their travel behaviour. There is evidence from Perth, WA that few people are aware of the options offered by public transport, and not many more are aware of bicycling and walking options (Socialdata, 2000). Voluntary behaviour change may be defined as *change that occurs when individuals make choices for personal reward without a top-down mechanism, regulation of any sort or a feeling of external compulsion ... an individual decides to make a change so that he or she will improve their personal life in some way* (Ampt, 2003, p. 4).

One approach (IndiMark®) is based on informing people about the transport-related infrastructure and services that are available in the locality. The emphasis is generally on improving the practical understanding of the transport system and services that is held by individuals. Where this understanding is incomplete, this improved knowledge and human capacity leads to shifting travel from environmentally unfriendly modes (principally driving the car, and especially driving alone) to environmentally friendly modes (public transport, walking, bicycling, or riding in the car as a passenger), without significantly altering the number of out-of-home activities in which people engage, and also probably not changing the amount of time spent travelling. Indeed, in this approach, it is often seen as most successful when there is no change in the number of trips, but a reduction in VKT by car as driver, in particular from before the application of the approach to after.

The other approach (Living Change/Living Neighbourhoods®) focuses on the problems that individual households have in travel and activities in their community. It describes problems in terms of their personal effects on the individual and households, and provides opportunities for individuals and households to solve their own problems. In so doing, members of the community also assist in solving community or government problems. This approach is as likely to lead to an overall reduction in travel as to a shift in mode. However, in Living Change or Living Neighbourhoods®, success would be considered to have been achieved if overall person kilometres of travel (PKT) were reduced, and possibly the number of trips and out-of-home activities were reduced, without necessarily changing the travel mode at all. Rather, this approach seeks to offer possibilities to reduce travel overall, albeit with the intent that travel by car, especially drive alone, will be reduced the most.

Thus, the travel behaviour modification sought is a reduction in the use of drive alone, especially, either through a shift of some car driver trips into more environmentally-friendly modes without changing overall levels of travel, or through blending of activities, destinations, modes, or over time, to reduce the overall need for travel. The outcome in either case would be a reduction in the level of car traffic, without the imposition of politically unpleasant strategies, such as pricing or tolling options, and without the need for significant investment in transport infrastructure. Indeed, much of the emphasis of these approaches is towards making better use of the currently-available infrastructure, and particularly using unused capacity, whether by mode, or by time of day.

The appeal of these strategies is obvious. Providing improved information is much cheaper than investing in new infrastructure. If such informational enhancement can lead to significant shifts in travel, whether reducing overall travel or shifting travel out of the car, it offers the potential to alleviate some of the congestion that seems to have become ubiquitous in early 21<sup>st</sup> century cities, without incurring major capital outlays. It also can contribute in a substantial and significant way to reduction of greenhouse gases, through reductions in vehicle kilometres of travel (VKT) involving cars, which are the major contributors to the production of CO<sub>2</sub> in most developed countries of the world.

As a result, the necessity of evaluation is clear. While these policies are much cheaper than infrastructure investments, they still come at a price, and investing in them needs to be determined as being worthwhile or not. Comments have been heard that it is strange that these policies should be subjected to a demand for such an intensive evaluation, when many other transport policies are either never evaluated or are evaluated only very superficially. There are probably three reasons for an intense scrutiny of this approach, at least. First, it generally sounds to transport planners to be too good to be true. Transport planners are used to having

#### *4 Evaluating voluntary travel behaviour interventions*

to spend large sums of money (in the hundreds of millions of dollars, if not in the billions of dollars) to achieve very modest mode shifts, and often achieve much less than was expected or hoped for. Second, because the approach may cost hundreds of thousands of dollars to implement in large areas, there is an unwillingness to invest even these modest (by comparison with infrastructure projects) amounts of money in a policy that may possibly not produce the claimed benefits. Third, transport planners find themselves facing a situation in which no option that has been tried in the past seems to work. They have accepted to a greater or lesser degree that we cannot build our way out of congestion. At the same time, initiatives to get people to ride public transport through new investments in public transport, especially rail investments, simply fail to deliver the shifts from car use that were hoped for and expected (Stopher 2004). Therefore, when an approach appears that seems possibly able to provide a policy option that might actually work, especially when that approach is so different from what the mindset of planners has been in the past, there is naturally scepticism that the approach can work. Hence, intense scrutiny of these approaches is almost inevitable.

#### **Evaluating Voluntary Travel Behaviour Change**

The most difficult aspect of voluntary travel behaviour change is evaluating the extent to which it results in actual travel behaviour changes, and assessing the sustainability of those changes over time. Assessing changes in travel behaviour is not isolated to voluntary travel behaviour change. However, it has taken on prominence because of the possibility that significant and substantial changes in car driver use and vehicle kilometres of travel can be achieved with what is comparatively a very low cost process. If it were, indeed, true that shifts of 10 percent or so of car driver trips to other modes of travel can be achieved by this mechanism, then the cost-benefit ratio of VTBM is extremely attractive and would warrant extensive adoption of this approach for such goals as reducing enhanced greenhouse gas emissions and delaying more capital-intensive investments in the transport system. For example, there is a National Travel Behaviour Change Program in Australia which is looking for an outcome from voluntary travel behaviour change that will have over 186,000 households participating in the Australian Capital Territory, Queensland, South Australia, and Victoria. Over a five-year period, it is projected that these households will reduce vehicle kilometres of travel by more than 3,850 million, and will thereby reduce carbon dioxide emissions by over 1.2 million tonnes for that period (Pramberg, 2004). Further, if these are sustainable changes, then the reductions would actually continue for many years after that. If such reductions can be obtained through this mechanism, then it is indeed a very powerful one.

There are several difficulties involved in evaluating this policy. First, there is a need to be able to measure the travel behaviour of participating households for a period before the intervention, sufficient to establish what are normal travel patterns, vehicle kilometres of travel undertaken, and mode use for those households. Second, there is a need to be able to measure travel patterns at various times after the intervention, to establish whether or not, and how much, travel patterns have changed, and whether or not the changes are sustained over the long term. Two difficulties arise immediately here, in that travel patterns are not static at the best of times. Therefore, there is a need to have control groups, that will allow determination of what are the probable changes that would have taken place in household travel without the intervention, over the same period of time. There is also the further problem that households themselves are not static, and that changes in household structure may take place that would result in additional changes in travel behaviour that will be confounded with

the changes caused by the intervention itself and by the passage of time and the effects of external factors.

Third, the amount of change that is expected is not large in statistical terms. To measure small changes accurately requires large sample sizes, which poses yet another difficulty. Finally, all of this measurement requires some quite burdensome activities by those who have agreed to participate in the VTBC Program. Motivating the response to measurement and maintaining it over a period of time is a substantial challenge.

### How to Measure Change

There are two primary options to consider for measuring a change in a population. One is to perform two or more surveys, using independent cross-sectional samples of the population for each survey. The other is to form a panel, and repeat measurements on the same sample on each occasion. There are pluses and minuses to each of these possibilities (Golob *et al.*, 1997).

*Repeated Cross-Sections:* The sample size required to measure a certain change with a given level of precision will generally be substantially larger for the repeated cross-sectional samples than for a panel (Richardson *et al.*, 2003; Kish, 1967). Therefore, each cross-sectional survey sample will need to be large. Generally, the difference in sample size between a panel and repeated cross-sections for a phenomenon such as travel is likely to be on the order of a factor of two or more. This results because the variance of the difference between two occasions is reduced by the covariance between the two occasions. If the two samples are independent, the covariance is normally assumed to be zero. In a panel, however, the covariance term can be large, and may be close to the sum of the variances on the two occasions, thereby reducing the variance in the difference substantially.

An example will serve to illustrate this. Suppose the voluntary travel behaviour change is being applied to a reasonably large population of say 30,000 households. Suppose that it is expected that a change of 10 percent will occur in car driver kilometres over the entire population of 30,000 households. Suppose, further that it is desired to estimate this change with 95 percent confidence to within  $\pm 2$  percent, i.e., that we would have 95 percent confidence that the actual change lies somewhere between 8 percent and 12 percent. Suppose that the mean of car kilometres of travel per household is around 40 per day, and that the standard deviation is of the order of 48 kms per day (see Richardson *et al.*, 2003). The sampling error for the difference between two occasions is given by (Kish, 1967):

$$s.e.(\bar{x}_2 - \bar{x}_1) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where:

$\bar{x}_1, \bar{x}_2 =$  mean car kms of travel per household per day

$s_1, s_2 =$  standard deviations for the before and the after surveys for VKT per household per day

$n_1, n_2 =$  sample sizes on the two occasions – before and after

## 6 Evaluating voluntary travel behaviour interventions

In this example, the values of  $v_1$  and  $v_2$  would be 40 and 36; the values of  $s_1$  and  $s_2$  would be 48 on the first occasion and we will assume a small decrease to 43 for the second occasion. Assuming that the sampling error is to be  $\pm 2$  percent or  $\pm 0.8$  kms at 95 percent confidence, then the allowable sampling error is  $\pm 0.8$  kilometres divided by 1.96, or  $\pm 0.41$  kms. Assuming that we wish to draw the same size sample on each of the two occasions (before and after), then the above equation can be manipulated and rearranged to give the required sample size as:

$$n = \frac{s_1^2 + s_2^2}{(s.e.(\bar{x}_2 - \bar{x}_1))^2}$$

This comes to a sample size of 24,706 households. Even after applying the finite population correction factor, the sample size is still 13,548 households on each occasion.

On the other hand, cross-sectional samples are much easier to draw. Provided that the population from which the sample is to be drawn is large enough, the level of non-response is likely to be similar in each cross-sectional sample. Also, repeated cross-sectional samples will reflect changing demography in the population, changes in household composition, and other similar changes that may be taking place over the period of the surveys.

*Panel Surveys:* These have much lower sample size requirements, although the above scenario still generates a large sample size requirement. In the case of a true panel (i.e., where all households measured on the first occasion are again measured on the second occasion), the sampling error is given by:

$$s.e.(\bar{x}_2 - \bar{x}_1) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} - \frac{2 \text{cov}(x_1, x_2)}{n_1}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} - \frac{2R_{12}s_1s_2}{n_1}}$$

where  $R_{12}$  is the correlation between the two occasions. If we use the previous example and assume that this correlation is 0.9, meaning that there is a lot of similarity between the two occasions, then the sample size required is given by:

$$n = \frac{s_1^2 + s_2^2 - 2R_{12}s_1s_2}{(s.e.(\bar{x}_2 - \bar{x}_1))^2}$$

This gives a required sample size of 2,604 households. If two independent cross-sectional samples used this sample size, then the sampling error would be  $\pm 1.26$  kms. At 95 percent, the confidence range would be that the change was equal to 4 kilometres plus or minus 2.48 kilometres, or an error level of  $\pm 6.2$  percent. This is more than three times the desired level of accuracy, which was specified as  $\pm 2$  percent.

Apart from substantially lower sample size requirements, the other principal value of the panel survey is that it measures the dynamics of change much more clearly than repeated cross-sectional samples. In a panel, comparisons can be made for the same household on each occasion that measurement is undertaken, and the magnitude of change is directly measurable. At the same time, it is also possible to track the magnitude of other changes that may have taken place in the characteristics of the household. In repeated cross-sectional surveys, it is only possible to make aggregate comparisons, and there is an underlying assumption that the households in one cross-sectional survey are equally representative of the population as those

in the next survey. Even if the samples are stratified into certain socio-demographic groups, comparisons are still at an aggregate level, and the errors are not significantly reduced. Uncovering underlying causes for dynamic changes is generally not possible in repeated cross-sectional surveys.

For example, repeated cross-sectional surveys may show that bus ridership is remaining constant over a period of years. From this, it may be assumed that bus riders comprise a particular segment of the population, so that marketing efforts should be targeted to car users, to persuade them to shift from car to bus. However, a panel survey may reveal that, while the total number of bus riders remains unchanged, the individuals who are riding the bus are constantly changing. The panel survey might reveal that as young adults leave school and begin to work, they become bus riders. However, as they begin to amass savings and to establish an earnings level that permits it, they purchase and begin to use a car, in place of riding the bus. This would suggest a completely different marketing approach, in which the effort would be to retain bus riders, as they become more wealthy, and to persuade them not to stop using the bus.

There are a number of problems that arise with panel surveys, however. First, panels are subject to attrition. This is the loss of panel members through several mechanisms, including households that move away from the study area, households that cease to exist due to some form of break up, and households that drop out of the panel. In the U.S., for example, it has been estimated that panel attrition may run as high as 30 percent per year, for a panel conducted on an annual basis. More frequent measurement of the panel may reduce attrition due to moving and break up, but will often increase attrition due to households becoming tired of participating. Another problem with panels is that of "conditioning". This is the process in which households involved in repetitive surveys either change their behaviour because of what they learn through participation, or become able to provide false answers that they deem to be more politically "correct" or desirable in response to the survey questions. In this process, the households become less and less representative of the population, because their participation in the panel is changing what the panel measures. At the same time that this occurs, there is also panel fatigue, which results in panel members giving less attention to the survey task in repeated waves, and eventually dropping out of the panel.

Another issue is that the panel may become less representative of the population as time goes by, depending on how attrition is handled. There are three principal methods available to deal with attrition. The first is to oversample the panel initially, so that the panel declines in size over time, due to attrition, and each successive wave of the panel measures a decreasing subsample of the original panel. The second is to make up for the attrition at each wave of the panel, by selecting new members of the population to replace the lost households, where the replacement households are chosen to be as similar as possible to those who have been lost from the panel. The third is to make up for the attrition by selecting replacement households that reflect, as far as possible, the changing nature of the population. In this method, efforts would be made to change the constitution of the panel at each wave, so that it remains representative of the current population, as far as panel make up will allow (Hensher, 1987).

To reduce attrition in panel surveys, it is also usually necessary to keep in contact with panel members between survey waves. This helps to keep the interest of the panel members and may also help in finding out new addresses and other changes as they occur. However, this adds to the expense of the panel survey, offsetting some of the gains that are made by not having to draw fresh samples for each survey. Overall, however, panels appear to offer

## 8 *Evaluating voluntary travel behaviour interventions*

sufficient potential gains over repeated cross-sections that they should be considered seriously as the potentially best method to measure travel behaviour change (Purvis and Ruiz, 2003).

### Issues of Variability

As mentioned earlier, there is a problem of the variability in travel from day to day, from season to season, and in response to external stimuli. Variability is highest at a person level, and considering daily travel without controlling for the day of the week (Richardson, *et al.*, 2003). If a panel survey is conducted, in which the same day of the week is used on each occasion, the variability decreases fairly significantly, dropping to about three-quarters of the uncontrolled day of week variability. If it were possible to measure travel for a whole week in each wave of the panel, the variability drops to a bit less than one half of the uncontrolled daily variability. However, there has been very little success in the past with gaining cooperation of respondents to complete a travel survey for a whole week. There are also known problems with multi-day surveys of a drop in reporting completeness with increasing length of the survey (Madre, 2003; Purvis and Ruiz, 2003). If a conventional travel survey approach were used, it would appear that the best option may be to use a two-day diary, and to control this so that, in a panel, each household is asked to complete the diary on the same two days of the week as in previous waves. Two-day diaries, although still prone to reporting drop off on the second day, will provide some indication of day-to-day variability, without resulting in excessive respondent burden, whilst the use of the same days of the week on each wave of a panel would contribute to some substantial reduction in variability and concomitant improvement in accuracy for a given sample size.

### Choosing a Control Group

This is another area of the evaluation that poses some considerable difficulties. The issue here is first and foremost how to select an appropriate control group. The control group needs to be matched to the target population in terms of socio-demographics, location with respect to major employment centres, levels of public transport service and orientation of those services, car ownership and use rates, and trip-making rates and patterns. At the same time, the control group needs to be sufficiently physically removed from the target group that there will not be diffusion of the VTBC program to the control group population. Given that most applications of VTBC have been to specific suburbs within a large metropolitan area, the selection of an appropriate control group is by no means trivial. At the same time, failure to select a control group appropriately may make it impossible to determine if there are underlying travel behaviour changes, resulting from other factors such as changes in fuel prices, that need to be factored into the changes measured within the target population. There is no simple solution to this problem. Rather, there is a need to use all means at one's disposal to select matching populations, considering such things as recent census information, journey-to-work travel information, geographic positioning, public transport service levels, and any other available information that is descriptive of the travel and locational situation, and to choose locations that are far enough away from the target areas to be immune from diffusion effects.

A related issue for the control group is motivation to participate in the repeated surveys. With the target group, there is at least the notion that there is some benefit being received by households that participate in the VTBC program. For the control group, there is no such benefit. As a result, response rates are likely to be lower from the control group, particularly

from repeated measurements, resulting in the potential for more serious biases to exist in the control group. It is probably necessary to consider providing incentives to such respondents to participate in the surveys and to continue to do so, in a panel survey.

### Measurement Error

A further problem that pervades all of the discussion on evaluating VTBC is that it is highly dependent on self-reporting. No matter whether households are interviewed through face-to-face surveys or over the telephone, or are provided with paper diaries to complete and return or an internet site to undertake the survey, all of these methods of survey involve self reporting, usually relying on some degree of memory of recent travel and activities undertaken. It has been known, since the first household travel surveys were conducted, that people under-report their travel. Regular and repeated trips are usually reported well (e.g., trips to work or to school). Infrequent trips, especially short trips, are often forgotten, or are completely misreported. Recent work in the U.S. has shown that, using GPS as a means to check on the completeness of diary reporting, the level of under-reporting may be as high as 20 percent of daily trip making, although the trips that are omitted are generally short trips (Wolf et al., 2003). Unfortunately, it may well be the short trips that are not reported that may be most susceptible to change under VTBC programs. At the same time, any omission of trips, which may be different among different individuals, jeopardises the measurement of the extent of the changes that take place.

Possibly the least problematic survey method, and one that may lend itself to much of the evaluation of the long term effects is the use of odometer surveys, in which responding households are asked to report the odometer readings of each of the household's vehicles at the beginning and end of a survey period. This will provide a means to determine the overall VKT driven by responding households, and would generally provide information about the reductions in VKT achieved. However, in order to determine whether there are biases in the sample, and to have a mechanism for expanding the sample, various household and person characteristics will also need to be collected each time that odometer readings are obtained. There is also no knowledge as to how accurately people report odometer readings in a telephone, written, or internet survey. We do not know if they go and look at each car's odometer, or give a guess, based on the last value they recall seeing. A face-to-face interview, with an interviewer visit to the home for both the start and end of the survey period, would be the only method in which there would be reasonable certainty about the veracity of the reported odometer readings. This method will not permit assessment of shifts to public transport, walking and bicycling, nor will it permit determination of whether travel has been reduced overall. It will also not provide indications of changes in vehicle occupancy, or of trip chaining.

### External Evidence

One of the other measurement devices that should be considered is that of external evidence of travel behaviour change. By this, we mean evidence that is not gathered directly from households or persons who may have modified their behaviour, but rather from external evidence, such as bus ridership, public transport ticket sales, and measurements of road volumes. Changes in bus ridership and ticket sales will again require lengthy periods of before data to establish trends, and also to cover areas that are not expected to be affected by

the VTBC initiatives. This would allow more reliable estimation of whether there have been positive changes in the areas of application of VTBC in public transport use that may be attributable to the travel behaviour changes. In the case of spot counts of road volumes, it is likely to be a little harder to establish, because the relative change in car driver use is much smaller than that for public transport use. (For example, if public transport is carrying four percent of travel in a region, and car accounts for 64 percent, a shift of five percent of car trips into environmentally friendly modes is rather small. However, if two-thirds of those trips find their way onto public transport, the local share of public transport may increase by as much as 15 or 20 percent, which is much more readily evident and measurable.) Corroborative evidence from public transport use and from road counts would be an important addition to the reliability of any measurement undertaken on the persons and households themselves.

### Diffusion Effects

Almost all of the evaluations to date have concentrated on applications of VTBC programs to households and communities. Measuring the diffusion effects of these applications has not been attempted. In addition, evaluating the diffusion effects of program applications in schools, workplaces, and other special locations has not been attempted either. The major difficulty here is in identifying the population to be measured. For workplaces, this may not be too difficult, since workers at many employment locations will not change substantially over a period such as five years. However, populations of schools and universities are much more transient, and there are likely to be much more significant issues about how to measure these populations to determine if there is ongoing travel behaviour change.

Further, it is obviously the hope that such applications as these will be further diffused into the community through word of mouth, and reinforcement from “good stories” about changes that people have made. Given this, the measurement of the diffusion effects becomes challenging indeed. First, it would appear to be necessary to measure travel behaviour in those communities from which workers, schoolchildren, or university students are drawn. Second, to determine if the VTBC programs at those workplaces, schools, and universities have had an effect in the community, it becomes necessary to ask if respondents have heard about TravelSmart, or any other brand name used for the VTBC programs. This almost immediately mandates that a panel cannot be used in this application. Asking this question of members of a panel will sensitise them to the existence of the program and is more likely than not to cause subsequent behaviour changes. Therefore, a program of non-overlapping cross-sectional surveys appears almost mandatory for this assessment. As noted before, however, this will have serious impacts on the sample sizes required and therefore on the expense of the survey.

One of the strategies that is likely to be necessary to measure the diffusion effects is to find suitable control groups. However, because it is unlikely that control groups can be found in reasonable proximity to the target populations, it will be necessary to begin measurement of travel behaviour in both the potential target area and the control area at least two years before the VTBC program is to be initiated. What this will achieve is the establishment of trends in both the target and control populations, prior to any intervention. Then, as trends in the control group are observed after intervention with the target group, it becomes possible to estimate the probable changes in travel behaviour that would have occurred in the target group, without the intervention. Hence, the actual diffusion effects can be estimated.

## Focus Groups

Focus groups are often useful to determine qualitative information that can answer certain types of questions, such as questions about what strategies within a TravelSmart initiative worked and which did not for those individuals who participate in the focus group. Focus groups can also be used to help design a good measurement instrument for the changes in travel behaviour. However, focus groups, in and of themselves, cannot answer the questions that are desired from an evaluation, which are: how much change in travel behaviour has taken place, and what could one expect to be the impact if applied to a larger population. By their nature, focus groups neither provide reliable quantitative information nor provide information that can be expanded to a population. Therefore, they do not provide an answer to the primary need to evaluate the size of the travel behaviour change.

## Evaluation of TravelSmart “Households on the Move”

The ACT commissioned a TravelSmart pilot called “Households on the Move”. It aims to target households at a ‘change moment’, that is when they are moving house, and are potentially open to changes in travel habits, as they accustom themselves to a new environment for travel and activities. The project is described in more detail elsewhere (Ampt et al., 2004). This project also included an evaluation component.

Two principal groups of households were originally targeted for the intervention: households that were about to move (called the *pre-move* households), and households that had recently (within the past 30-45 days) moved (called the *moved* households). The aim was to undertake the intervention with one hundred households. Of these, it was desired to have as many as possible participate in the evaluation. For the pre-move households, no before survey is possible, because they are provided with travel behaviour change tools before they have moved, and will implement them in their new home location. Therefore, how they are travelling before the move will have no bearing on measuring how they are travelling after the move. It was planned to measure the travel behaviour of these households only after they had moved, and to measure, as a control group, pre-move households that were not interested in the intervention tools. For the moved households, the control group again needs to be similar households. In this case, because of the pilot nature of the project, it was decided to obtain the control group again as those households that were offered the travel behaviour change tools, but chose not to accept them. For the moved households, a before and an after survey were planned. In this case, the differences between before and after the intervention should give an indication of the extent to which households actually changed behaviour as a result of the travel behaviour change tools.

Two methods were designed to measure household travel behaviour. The core method was the use of GPS devices, both wearable ones and in-vehicle ones. This was supplemented by a two-day diary survey<sup>1</sup>. The GPS devices are passive devices (Stopher *et al.*, 2002), requiring almost no intervention on the part of the respondent. The in-vehicle devices are simply plugged into the accessory/cigarette lighter outlet, and the antenna/receiver is placed on the vehicle dashboard. A bag, containing the recording unit, can be placed in any convenient location, such as under the front passenger seat. The wearable device comes as a small bag, with a shoulder strap. The antenna/receiver is mounted on the top of the shoulder strap, while

---

<sup>1</sup> At the time of writing, the before survey is almost complete. The after survey is due to commence in several weeks, but results from this should be partially available for the final paper for the ATRF meeting.

## 12 *Evaluating voluntary travel behaviour interventions*

the recording device and a small battery pack are contained in the bag. The entire unit weighs about 500 grams. The wearable device needs to be plugged into a power supply overnight, to keep the battery pack charged, but requires no other action, except to wear it.

Each household was approached first to determine their interest in the travel behaviour change tools. Households were then asked if they would be willing to be involved in the GPS survey. If the household indicated that they would not be willing to undertake the GPS survey, they were asked if they would complete a travel diary for two days, with an expected follow up of another diary in a period of two to four months. The target group was that group of households that were interested in the travel behaviour change tools, and were willing to complete either the GPS survey or the diary survey. The control group was those households that were uninterested in the travel behaviour change tools, but were willing to either engage in the GPS survey or the diary survey.

The number of households that were approached was 294 of the moved households. Of these, 102 households agreed to receive tools. Of the remaining 192 households, 32 households agreed to participate in one or other of the two surveys, with 21 agreeing to do the GPS survey, and 11 agreeing to do travel diaries. However, 4 of the households for the GPS dropped out, and 5 households did not return travel diaries. Of the 102 households that agreed to the intervention, 26 agreed to do the GPS survey and 26 agreed to do travel diaries. However, 7 dropped out of the GPS survey and 16 did not return travel diaries. This left a total of 23 households in the control group (17 GPS and 6 diary), and 29 households in the target group (19 GPS and 10 diary). One of the main problems reported with the GPS survey was the size of the wearable GPS device, which several respondents found to be too large and obtrusive to wear to a new job.

Overall, the completion rate to both the GPS and the diary surveys was lower than desired and appear to be impacted a great deal by the fact that people who have just moved are not very willing to undertake these tasks at this particular time. However, the response rate to the GPS is higher than that to the diaries. For the control group, where an incentive of \$10 per household was offered to induce participation in the GPS survey, 11 percent agreed to do the GPS and 81 percent of those completed the GPS task. For the diaries, with no incentive, only 6 percent agreed to take the diaries, and only 55 percent of those actually completed the task. Similarly, for the target group, where no incentive was offered, 25 percent agreed to take the GPS devices, and the same proportion, 25 percent, also agreed to take diaries. However, 73 percent completed the GPS survey, whilst only 43 percent completed the diaries.

A further contrast is that the GPS survey collected a week of travel, whilst the diaries collected only two days of travel. The much richer data from the GPS should provide far better indicators of changes in travel patterns than can be obtained from the diaries. For example, we have conducted preliminary analyses, using twelve person records, and have also analysed the 16 households that returned diaries. Some statistics for these are shown in Table 1. We restricted the diary data to the first day only, because there is a drop off in reporting on the second day.

**Table 1: Comparison of GPS and Diary Person Travel Statistics**

Statistic	GPS Survey	Diary Survey
Mean VKT per person per day	37.25	50.30
Variance in VKT per person per day	1026.19	5921.15
Mean number of person trips per day	5.38	4.21
Variance in number of person trips	15.48	5.18
Mean VKT per person per weekday	40.71	50.30
Variance of weekday per person VKT	933.46	5921.15
Mean VKT per person per weekend	28.6	--
Variance of weekend VKT per person	1199.25	--
Mean weekday trips per person per day	5.75	4.21
Variance of weekday trips per person per day	13.38	5.18
Mean weekend trips per person per day	4.46	--
Variance of weekend trips per person per day	20.26	--

The kilometres per person came from respondent estimates of distance provided in the diary. This value is much higher than the actual VKT per day, given that people are known to incorrectly record distances travelled. On the other hand, it is noteworthy that the average number of trips per person per day is significantly higher from the GPS than from the diaries. In addition, from even a small number of observations, we are able to obtain estimates of weekday and weekend travel attributes, which are not possible from the diary data, where the number of observations is too small for the weekend data, although more households have been analysed. Another important finding is that the variance of the weekly data, as expected, is much smaller than the variance of the one-day data from the diaries (only the first day was used in this comparison).

## Conclusions

The task of evaluating TravelSmart initiatives is far from a simple one, and may be almost as expensive as the implementation itself, depending on the level of accuracy desired in measuring change resulting from the initiatives. Many difficulties arise in the evaluation process, including the sampling issues and also identification and measurement of a suitable control group. In terms of the future, there are several directions that become apparent as potentially useful ones that may help to solve some of these problems.

First, the use of GPS as a mechanism for measuring travel before and after the interventions holds out considerable promise. Even in the case of households that have just moved, the GPS devices appear to have been accepted more readily than standard diaries, and are able to produce data for a week's worth of travel, rather than the one or two days that are normally all that can be collected from a diary. As new technologies in GPS are developed (and there are much more miniature versions of the wearable devices now becoming available), the potential for gaining a higher acceptance of these devices is substantial.

Second, it seems clear that the advantages of a panel are so significant in this area that panels should be considered as the only appropriate method for evaluating TravelSmart initiatives. The enormous sample sizes that are needed to measure reliably the level of changes expected from a TravelSmart initiative are simply too small to be measured adequately by independent cross-sectional samples.

Third, external evidence of the changes in travel behaviour must be scrutinised carefully. Generally, it is necessary to have at least two full years of data available from before the TravelSmart intervention to be able to ascertain trends that were established before the intervention. Coverage of the control areas is also important, to be able to determine if there are other external effects in progress that might lead to different conclusions from the external evidence. Finally, it will become increasingly important to investigate the non-transport impacts of voluntary travel behaviour change interventions.

## References

Ampt, E. (2001) The Evaluation of Travel Behaviour Change Methods – A Significant Challenge, *Proceedings of the 24<sup>th</sup> Annual Australasian Transport Research Forum*, Hobart, April.

Ampt, E.S. (2003). ‘Voluntary Household Travel Behaviour Change – Theory and Practice’. *Proc 10<sup>th</sup> International Conference on Travel Behaviour Research*: Lucerne, Switzerland. <http://www.ivt.baum.ethz.ch/allgemein/pdf/ampt.pdf>

Ampt, E.S., P.R. Stopher, and R. Alsnih (2004). “Households on the Move – A New Application of Voluntary Travel Behaviour Modification”, paper to be presented at the 27<sup>th</sup> Australasian Transport Research Forum, Adelaide, S.A., September-October.

Ashton-Graham, C. (2004). Personal communication April 15, 2004.

Golob, T.F., R. Kitamura, and L. Long (editors) (1997). *Panels for Transportation Planning: Methods and Applications*: Kluwer Academic Publishers, Boston, 387 pages.

Hensher, D.A. (1987) "Issues in the Pre-Analysis of Panel Surveys", *Transportation Research*, (Special Issue on Longitudinal Methods, edited by D. Hensher), Vol. 21A, No. 4/5; 265-284. (Also Introduction: 247-248).

James, B. (1998), “Changing Travel Behaviour Through Individualised Marketing: Application and Lessons from South Perth”, *Papers of the Australasian Transport Research Forum*, Volume 22, Part 2, pp. 635-647.

James, B. (2002), “TravelSmart – Large-Scale Cost-Effective Mobility Management. Experiences from Perth, Western Australia”, *Municipal Engineer*, Vol. 151, Issue 1, pp 39-48.

James, B., W. Brög, E. Erl, and S. Funke (1999), “Behaviour Change Sustainability from Individualised Marketing”, *Papers of the Australasian Transport Research Forum*, Volume 23, Part 2, pp. 549-562.

Ker, I. (2004), “Household-Based Voluntary Travel Behaviour Change: Aspirations, Achievements, and Assessment”, paper prepared for the Transport Journal of the Institute of Engineers Australia.

Kish, L. (1967). *Survey Sampling*, John Wiley and Sons, New York, pp. 469-477.

Madre, J.-L. (2003). "Multi-Day and Multi-Period Data", in *Transport Survey Quality and Innovation*, P. Stopher and P. Jones (editors), Pergamon Press, Oxford, pp. 253-269.

Perkins, A. (2001). *The Greenhouse Abatement Potential of Travel Behaviour Change Initiatives*. National Taskforce on National Greenhouse Strategy Measure 5.3: Promoting Best Practice in Transport and Land Use Planning, Queensland Transport, Brisbane, Queensland.

Pramberg, P. (2004). *National travel Behaviour Change Program (NTBCP): Consultancy Brief and Offer Specifications: Long Term Monitoring*, Queensland Transport, Version 4, March.

Purvis, C.L. and T. Ruiz (2003). "Standards and Practice for Multi-Day and Multi-Period Surveys", in *Transport Survey Quality and Innovation*, P. Stopher and P. Jones (editors), Pergamon Press, Oxford, pp. 271-282.

Richardson, A.J., R.K. Seethaler and P.L. Harbutt (2003). "Design Issues for Before and After Surveys of Travel Behaviour Change", Proceedings of the 26<sup>th</sup> Australasian Transport Research Forum, Wellington, New Zealand, October, 22 pp. On CD-ROM.

Socialdata (2000). *Potential Analysis "Perth"*. Report to the Department of Planning, Western Australia, <http://www.dpi.wa.gov.au/travelsmart/pdfs/Report.PDF> , accessed 16<sup>th</sup> April, 2004.

Stopher, P.R., P. Bullock and F. Horst (2002). "Exploring the Use of Passive GPS Devices to Measure Travel," in *Proceedings of the 7<sup>th</sup> International Conference on Applications of Advanced Technologies to Transportation*, edited by K.C.P. Wang, S. Medanat, S. Nambisan, and G. Spring, ASCE, Reston VA, pp. 959-967.

Stopher, P.R. (2004). "Reducing Road Congestion: A Reality Check", *Transport Policy*, Vol. 11, No. 2, pp. 117-131.

Wolf, J., M. Loechl, M. Thompson and C. Arce (2003). "Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys", in *Transport Survey Quality and Innovation*, P. Stopher and P. Jones (editors), Pergamon Press, Oxford, pp. 483-498.