

PERSONALISED MARKETING – IMPROVING EVALUATION¹

Dr Carolyn O'Fallon
Pinnacle Research

Dr Charles Sullivan
Capital Research

ABSTRACT

This paper provides guidance to local authorities and others who may wish to apply a personalised marketing programme in a given area and to measure the impact of such an intervention.

'Personalised Marketing' is used to describe a programme aimed at changing people's travel behaviour by a combination of education, persuasion and provision of personalised information to either individual households or individual people. One of the best-known personalised marketing programmes in Australasia and Europe is "IndiMark"® or "individualised marketing".

Various personalised marketing demonstration programmes have claimed substantial success in decreasing car use and increasing trips by alternative modes, thereby convincing some local authorities that such programmes may be the "panacea" to congestion problems in urban areas. Our recent involvement in helping to plan the evaluation of a personalised marketing trial in Birkenhead (Auckland, New Zealand) caused us to review the international experience with various trials and their evaluation.

Our investigation is two-pronged:

- ▶ To learn to what extent we can guide "pre-selection" of the area and participants / households for a personalised marketing initiative
- ▶ To learn from pitfalls in the evaluation of the impacts of such programmes.

The evaluation ("after") and elicitation surveys in Birkenhead highlighted the inadequacies of the current public transport system, providing evidence to support the widely-used pre-selection criteria of a "good quality" public transport service being in place prior to the initiation of a personalised marketing programme (Department for Transport 2002; TAPESTRY, 2001).

Analysis of the "before" survey data revealed characteristics distinguishing "receptive" from "non-receptive" individuals and households. Not only did the analysis reveal significant differences in personal characteristics and attitudes towards "environmentally friendly modes", but also we found indications of pre-existing differences in mode use, which could confound potential evaluations of mode change *after* the intervention (e.g., as claimed in some analyses of individualised marketing in South Perth).

¹ This is a revised version of an earlier paper presented at the 26th Australasian Transport Research Forum, 1-3 October 2003, Wellington, New Zealand.

Furthermore, great care must be taken with respect to the sample used for the evaluation. The statistical power to detect *significant* differences between before- and after-measurements is determined not only by sample size but also by the variability of behaviour. If people vary greatly in the number of trips driven and/or distance driven on a day-to-day or week-to-week basis (even in the absence of an intervention), then larger sample sizes (and/or longer data collection periods than the usual one-day trip diary) are needed. We used the 1997/98 New Zealand Travel Survey to estimate day-to-day variability in transport behaviour (both distance and trips, for several different modes). Results suggest that the sample sizes required are distinctly larger than seen in some recent research locally and that traditional travel diary methods may not be cost-effective for impact evaluation of many small pilot projects.

1. INTRODUCTION

1.1. BACKGROUND

Traditionally, decision makers have had three approaches to encourage changes in traffic:

- (1) modify the supply of networks (build more capacity)
- (2) modify land use
- (3) manage the demand through various means (travel demand management techniques).

Many governments and local authorities have now come to the conclusion that it is no longer acceptable to simply build more roads to meet increased demand. Modifying land use is more of a medium- to long-term policy option, while travel demand management is potentially a “quick fix”.

Voluntary, and in some cases mandatory, travel behaviour change programmes are an increasingly common part of the travel demand management tool kit. There are several options. Some target whole populations or specific groups within a population – for instance, travel awareness campaigns, employer travel plans, rideshare or carpooling programmes, or school travel plans. Others are personalised approaches such as journey planning, “Travel Blending”®, “IndiMark”®, or “TravelSmart”. Generally speaking, personalised marketing initiatives encourage participants, through a combination of education, persuasion and provision of personalised information, to take action to reduce their vehicle use in favour of more “environmentally friendly modes” such as walking, public transport, cycling and car pooling.

1.2. THE BIRKENHEAD DEMONSTRATION PROGRAMME

This analysis in this paper focuses on a personalised marketing demonstration programme carried out in Birkenhead, Auckland, New Zealand in 2002 was broadly based on the “IndiMark”® or “individualised marketing” programme, established and trademarked by Socialdata. Overviews of the “IndiMark”® process are found in several sources (see for example Brög and John, 2001; Marinelli and Roth, 2002; and UITP / Socialdata, 1998) and so will not be repeated here. The process for the Birkenhead demonstration programme can be described as follows:

Personalised Marketing – Improving Evaluation

- (1) Randomly selected households were sent a package containing a letter from North Shore City Council, an explanatory letter, household form and 2-day travel diaries (including questions about attitudes towards alternative modes and about interest in receiving further information on such modes) for up to four people over the age of 14 to complete.
- (2) Phone calls were made to encourage and offer assistance to complete the household forms and travel diaries.
- (3) Based on whether or not additional information on walking, cycling, or public transport was sought by one household member, households (with completed travel diaries) were classified into programme groups.
- (4) Programme households were sent another letter and information request form. The request form included an offer of a free 1-month bus or ferry pass along with the potential for “home visits” about public transport, walking or cycling. A courtesy phone call was made to let them know that the request form was “on its way”. A further phone call was made if the form was not returned in 10 days.
- (5) Information was sent out, including free ferry or bus tickets to infrequent PT users as requested, and free visits were carried out.
- (6) An evaluation (“after”) survey and elicitation interviews were carried out with the programme group in order to evaluate the impact of the demonstration.

It is worth observing that, by combining the introduction to the programme with a 2-day travel diary, there was probably considerable non-response bias: the majority of those households returning partially or fully completed forms and diaries (over 70% of responding households) were “interested” in receiving further information about PT, walking and/or cycling. This preponderance is evident later in classifying individuals and households into “receptive and “non-receptive” categories.

An overall description of the Birkenhead programme and an evaluation of its impact has been provided elsewhere (Booz Allen Hamilton, 2003). Unfortunately, for reasons explained later in this paper, although there are strong indications of “positive” modal shift (i.e. less car driving, more public transport and/or walking) by the programme group, the final sample size in the “after” survey (n=139 households) means that it is impossible to confidently state the size or extent of any such change directly related to the Birkenhead demonstration programme.

Our initial focus in analysing the Birkenhead demonstration programme was to try to identify any characteristics that might distinguish individuals or households who may be “receptive” or “non-receptive” to a personalised marketing approach. The selection of the individuals and households for analysis is described in the ensuing sections. We also analysed responses to the “after” survey and “elicitation” interviews for clues as to what location characteristics could facilitate the successful implementation of a personalised marketing programme.

2. PRE-SELECTION CRITERIA FOR PERSONALISED MARKETING INITIATIVES

2.1. SELECTING INDIVIDUALS AND HOUSEHOLDS

2.1.1. International experience

We recently reviewed international experience² to ascertain key factors (such as characteristics of households, individuals and/or locations) or market segments that would assist in selecting participants for future personalised marketing initiatives. It was anticipated that the use of such “pre-selection” criteria would increase the likelihood of a successful outcome for the proposed scheme.

Unfortunately, although there was a substantial amount of information available on previous experience with personalised marketing initiatives, much of this was focused describing the process for each trial or experience and measuring the behavioural change associated with the programme. Very little data is yet available to describe the characteristics of the people, households or locations that commend themselves to successful personalised marketing approaches. At best, it is possible to glean an indication of what is viewed to be important in selecting participants for a programme from the published literature, such as Goulias et al. (2002) where the “IndiMark” approach used in a large South Perth trial is discussed in some detail. The information collected from respondents in the initial contact telephone survey is outlined as:

- ▶ Perceptions of public transport, walking and cycling in the area
- ▶ Number of people older than 14 years in the household
- ▶ The potential for increasing use of environmentally friendly modes by each individual
- ▶ Willingness to receive further information
- ▶ Any comments provided by the respondent.

While these characteristics may be important, the UK DFT (Department for Transport, 2002) observes that, even for the personalised marketing approach – which is the best developed of all the types of alternative mode marketing initiatives (including “travel blending”, personalised journey planning, and travel awareness campaigns) – patterns are difficult to ascertain, the small number of cases makes statistical validation of any findings impossible, and that there is no conclusive pattern emerging as to when and where it is most useful. A recent article (Roth et al. 2003) asserts that “individualised marketing” is appropriate for a broad range of urban areas, but does not provide sufficient evidence to override the DFT’s conclusion.

The recently completed phase (TAPESTRY 2001) of the European Commission’s organised work-stream focused on “transport awareness campaigns” (incorporating same initiatives as the DFT study) also concludes that “segmentation” for all types of marketing initiatives is “in its infancy” and that socio-demographic or socio-economic

² This review is reported more fully in an unpublished document prepared for the Energy Efficiency and Conservation Authority (NZ).

criteria cannot yet be clearly identified, but are not the only means for segmentation. With respect to personalised marketing initiatives, the report suggests several alternative groupings, including targeting:

- ▶ those who want to change their transport behaviour due to a pre-existing favourable attitude
- ▶ those who already use public transport, or walk or cycle (to increase their use)
- ▶ those who are interested and want to take part
- ▶ the young (who have an ability to influence adult behaviour)
- ▶ car dependent *journeys* as opposed to car dependent *people*
- ▶ people at times of change in their lives, such as moving house, changing jobs, starting a family, etc.

With respect to location, the primary pre-selection criterion noted is that the area already has a good quality public transport (PT) service – and that there is a perceptual gap between the actual quality of the service and what people *believe* exists. Indeed, DFT (2002) concludes that:

“It is clear that the techniques will generally only work ‘on their own’ where there is a large gap in perception between what exists and what people believe exists. For public transport, where services and travel quality is much higher than is perceived, personalised approaches can have very large effects, but where such a gap does not exist the travel behaviour effects could be negligible.”

They hold that this is true for any environmentally friendly mode – PT, walking or cycling.

With this information in hand, we turn to the Birkenhead demonstration programme to see what may be learnt about the individual, household and area characteristics.

2.2. WHO IS RECEPTIVE TO PERSONALISED MARKETING? – BIRKENHEAD INDIVIDUAL ANALYSIS

Birkenhead is a northern suburb in Auckland (formally, within North Shore City). It is separated from the Auckland CBD by a major harbour bridge that has been severely congested in recent years. Public transport services include a ferry service across the harbour, in addition to bus services.

2.2.1. Selection of individuals for analysis

In the 2002 Birkenhead trial, the programme group selected for potential delivery of personalised marketing information were only those who:

- ▶ answered “yes” to either one of the statements “I would like to receive some information on bus and ferry services in Birkenhead” or “I would like to receive some information about walking and cycling in Birkenhead and greater Auckland”
- ▶ and completed a 2-day travel diary.

Those who answered “yes” to the statement “I am interested in receiving a free pass to try out the bus or ferry”, but answered no to the other two statements were excluded from the programme group and thus had no further involvement in the trial.

Because our purpose is to identify individuals who may be receptive to a personalised marketing approach, we have constructed our groups of individuals somewhat differently. We have included in the “receptive” group, anyone who said “yes” to one or more of the three above-mentioned statements, irrespective of whether or not the 2-day travel diary was completed. However, if a regular public transport (PT) user responded “yes” to “I am interested in receiving a free pass” and “no” to the other two questions, they were excluded. Regular PT use was considered to be 3 – 7 times per week (this was recorded for each adult in the household independent of trip diaries). There are three reasons for this exclusion:

- (1) In a large scale application of personalised marketing, it is unlikely that there will be a requirement for households to complete a full 2-day, or even 1-day travel diary, because other options would be available to measure the effects of the programme (e.g., PT fare records). This will likely result in much higher participation rates than occurred in the Birkenhead trial.
- (2) International experience with personalised marketing initiatives found that people are more likely to actually use an alternative mode if they first indicate a willingness to try it (Goulias et al. 2002) or if they have some familiarity with it (which would be provided by a free ticket to use the mode) (DFT 2002; Curtis and Headicar 1997).
- (3) People already using PT, but at a low level (i.e. fewer than 3 times per week) were included because both UITP/Socialdata (1998) and Curtis and Headicar (1997) identified this population segment as one which could be encouraged to take more trips by PT.

2.2.2. Findings

In total, there were 1332 individuals aged 15 or older about whom personal data was received in the “before” survey. For some of these, the information given was limited to that contained on the “household form” (gender, date of birth, employment status, car use at / for work, and personal travel patterns in general). After excluding those who did not answer the questions about receiving information along with those who said “yes” to a free bus or ferry pass, “no” to the other two questions, *and* were regular PT users, there were 1199 people in the sample (53% were women). Of these, 73% (n=873) answered yes to one of the three “are you interested” questions, leaving 27% (n=326) in the “non-receptive” group.

Personal characteristics

Although significant, personal characteristics such as gender, employment status and age do not appear as strongly influential in receptiveness to alternative transport modes as do other characteristics or attitudes.

Women (55%) were slightly more likely than men (45%) to be interested in receiving additional material or free passes, $\chi^2(1)=4.845, p=0.025$.

Table 1 Employment status by receptive and non-receptive group

		Total N=1199	Yes to free pass &/or info No to all three	
	Count		N=873	N=326
Employment status	Full time work	55.0%	53.7%	58.3%
	Part-time or casual work	14.7%	15.9%	11.3%
	Student (secondary/tertiary)	9.2%	11.0%	4.3%
	Home duties/ other	9.3%	9.3%	9.2%
	Retired	11.9%	10.1%	16.9%
Total		100.0%	100.0%	100.0%

Table 1 shows that receptive individuals were part-time employees or students (27% c/w 16%) whereas non-receptive individuals were more likely to be retired, $\chi^2(4)=25.080$, $p<0.001$. Complementing the employment status data, Table 2 shows that a greater share (45%) of receptive individuals was less than 44 years of age than in the non-receptive individuals, where only 30% were under 44. A larger proportion of non-receptive individuals were older people (55+ years) than in the receptive group, $\chi^2(5)=24.47$, $p<0.001$.

Table 2 Age by receptive and non-receptive group

		Total N=1149	Yes to free pass &/or info No to all three	
	Count		N=836	N=313
Age in 6 categories	15-24	9.0%	10.2%	5.8%
	25-34	11.5%	12.4%	8.9%
	35-44	20.2%	22.1%	15.0%
	45-54	20.4%	20.0%	21.4%
	55-64	23.5%	21.3%	29.4%
	65 and over	15.5%	14.0%	19.5%
Total		100.0%	100.0%	100.0%

Receptive people were more likely to be in larger households (three or more people) than non-receptive individuals, who were more likely to be in 1 – 2 person households, $\chi^2(2)=9.565$, $p=0.008$. Larger households were also far more likely to have children in them – only 9 of the 237 households (4%) with 1 – 2 people had a child living there, compared with 145 of 227 households (64%) with 3 or more people. However, there is some evidence to suggest that although such households were interested in receiving information, once they had it, they found it too difficult to make changes:

“It made me more aware of the options available to me but because of my circumstances (kids), public transport is not an option, as it takes too long and is too expensive”

“It made me think about using the bus but because I have a family, it wasn’t convenient to take the whole family on the bus whenever we went out.”

Table 3 Household sizes by receptive and non-receptive individuals

	Count	Yes to free pass		
		Total N=1199	&/or info N=873	No to all three N=326
Number of people usually living at this address (compressed)	1	7.6%	7.1%	8.9%
	2	38.2%	36.0%	44.2%
	3 or more	54.2%	56.9%	46.9%
Total		100.0%	100.0%	100.0%

Higher receptivity of larger households may occur simply because more people respond to the survey, creating a greater probability that someone in the house could be interested in participating in the programme.

Travel behaviour

On the “household form” respondents were asked to indicate the frequency of use of various modes by each adult in the household. Receptive individuals were more likely to be infrequent or frequent PT users (using PT 1–7 days per week) those who did not request any information (18% c/w 7%) $\chi^2 (2)=25.110, p<0.001^3$. Table 4 also shows that a higher proportion (25%) of receptive individuals drive their cars less frequently (seldom / never / 1–2 days per week) than in the non-receptive group (17%), $\chi^2 (2)=7.325, p=0.026$.

Table 4 Frequency of PT use and car driving by receptive and non-receptive individuals

	Count	Yes to free pass		
		Total N=1199	&/or info N=873	No to all three N=326
Number of days per week using PT	3-7 days a week	9.3%	11.3%	4.0%
	1-2 days a week	6.0%	7.2%	2.8%
	Seldom / never	84.7%	81.4%	93.3%
Total		100.0%	100.0%	100.0%
Number of days per week driving a car	3-7 days a week	77.4%	75.4%	82.8%
	1-2 days a week	7.3%	8.0%	5.5%
	Seldom / never	15.3%	16.6%	11.7%
Total		100.0%	100.0%	100.0%

Interestingly, given earlier results in a New Zealand-based morning car commuter survey⁴, we found that the need to sometimes drive a car on employer’s business during the working day or driving a “company” car was not a statistically significant determinant

³ It may be useful to recall here that frequent PT users who requested a free pass but no further information are not included in this sample.

⁴ Our survey (O’Fallon et al., in press) of morning car commuters in Auckland, Wellington and Christchurch found that car drivers who drove company cars, used their vehicles for business during the working day 3 or more times per week, and/or who had car parks provided by their employers, were far less likely to consider alternative modes of travel for their “journey to work”.

of peoples’ interest in receiving information on other transport modes or the offer of free PT passes. However, it was clearly an issue for some respondents: “going to work I have a company car and a car park to park in so I prefer to drive to work”.

Vehicle availability

Table 5 shows that the ratio of vehicles to adults in the household is relevant to their receptiveness. The difference between the “receptive” and “non-receptive” individuals is reasonably minor albeit significant, $\chi^2(2)=7.489, p=0.024$.

Table 5 Ratio of vehicles to adults per household by receptive and non-receptive group

	Count	Total N=1199	Yes to free pass &/or info	
			N=873	No to all three N=326
Ratio of vehicles to adults per household	0-0.5 vehicles/adult	22.3%	24.2%	17.2%
	0.51 – 0.99 v/a	14.8%	13.9%	17.2%
	1 or more v/a	63.0%	62.0%	65.6%
Total		100.0%	100.0%	100.0%

Distance to nearest bus stop / ferry terminal

Intensive examination of the distances from home to the nearest bus-stop or to the Birkenhead ferry terminal failed to reveal anything statistically significant in terms of receptiveness or non-receptiveness of people to a personalised marketing initiative.

Attitudes

We collected responses to 3 attitude statements, rated on a 5-point scale (from *strongly agree* to *strongly disagree*), in the before survey:

- ▶ “Even if public transport (buses / ferries / trains) was free, I wouldn’t use it”
- ▶ “I’d ride a bike at least once a week if I had a good bike and I knew there was a safe route”
- ▶ “I like to walk at least once a day if possible”.

One additional statement offered respondents several options, namely:

- ▶ “I’d walk more often if...
 - There were better/more footpaths
 - There were better pedestrian crossings
 - There was less traffic on the roads
 - The traffic moved slower
 - Other (please specify)
 - Nothing could induce me to walk any more than I do now.”

Table 6 summarises responses to the first three attitude statements. All four of the statements showed statistically significant differences between receptive and non-receptive individuals, (χ^2 tests, $p<0.001$ in all three cases). The greatest variation occurred in the responses to the “free public transport” statement: 64% of those

requesting information / passes disagreed with this statement compared with 31% of those who did not request anything. The “bike” statement also had substantial differences: 33 % of those requesting information / passes agreed or strongly agreed with the statement compared with only 12% of those who did not request further information. The contrast between those requesting information and those not doing so was less for the walking attitude statement (73% c/w 59%).

Table 6 Responses to three attitude statements by receptive and non-receptive group

		Total	Yes to free pass &/or info	No to all three
<i>Count</i>		<i>N=1198</i>	<i>N=873</i>	<i>N=325</i>
Even if PT was free, I wouldn't use it	Agree or strongly agree	25.4%	20.3%	39.1%
	Neither agree or disagree	19.9%	16.3%	29.8%
	Disagree or strongly disagree	54.7%	63.5%	31.1%
Total		100.0%	100.0%	100.0%
<i>Count</i>		<i>N=1195</i>	<i>N=871</i>	<i>N=324</i>
I like to walk at least once a day if possible	Agree or strongly agree	69.1%	72.9%	59.0%
	Neither agree or disagree	19.7%	18.6%	22.8%
	Disagree or strongly disagree	11.1%	8.5%	18.2%
Total		100.0%	100.0%	100.0%
<i>Count</i>		<i>N=1193</i>	<i>N=870</i>	<i>N=323</i>
I'd ride my bike at least once a week...	Agree or strongly agree	27.4%	33.0%	12.4%
	Neither agree or disagree	21.8%	22.3%	20.4%
	Disagree or strongly disagree	50.8%	44.7%	67.2%
Total		100.0%	100.0%	100.0%

The single most common response to the walking statement “I’d walk more often if...” was “nothing could induce me to walk any more than I do now” (given by 39% of all respondents). This was followed by the category “other” (33% before recoding) which, when recoded, created a several new categories, but particularly “[I’d walk more often if] I had more time” (11%), along with several others each chosen by 2-3% of respondents, including “there was better weather”, “I were fitter / healthier”, “it was less hilly”, “it was shorter [distance] to where I want to go”, “it was safer for me personally”, and “if I was interested or not lazy.” The smaller categories have been left as “other” in Table 7.

Receptive individuals were more likely to respond with suggestions indicating that improving the walking environment (better or more footpaths, lighting, or pedestrian crossings) or decreasing or slowing the road traffic down would encourage them to walk more. . Non-receptive individuals were more likely to state, “nothing could induce me to walk any more than I do now” (52% c/w 34%).

Table 7 Response to “I’d walk more often if...” statement by receptive and non-receptive group

		Total	Yes to free pass &/or info No to all three	
	Count	N=1182	N=865	N=317
I'd walk more often if.....	There were better / more footpaths/ lighting / crossings	17.4%	19.9%	10.7%
	There was less / slower traffic on the roads	10.5%	12.1%	6.0%
	Nothing could induce me to walk any more than i do now	38.8%	34.1%	51.7%
	I had more time	10.8%	11.0%	10.4%
	Other (specify)	22.4%	22.9%	21.1%
Total		100.0%	100.0%	100.0%

Conclusion

All three of the attitude statements and a number of other personal characteristics and/or behaviours showed statistically significant differences between receptive and non-receptive individuals. Unfortunately, none of them are sufficient to provide a simple and definitive “screening” tool, which delineates those who can be assigned to the “receptive” category over the “non-receptive” category. To use any one or more of the characteristics or other statements in such a way runs the risk of missing a number of people who would still be interested in obtaining information about PT services, walking and cycling in their area. It may be that, given a sufficiently large enough sample to measure the impact of a personalised marketing initiative, some of these variables could be found to delineate who will follow through on their interest to measurable change in behaviour, but this remains to be seen.

2.3. WHO FOLLOWS THROUGH ON THEIR INTEREST? – BIRKENHEAD HOUSEHOLD ANALYSIS

2.3.1. Selection of households for analysis

Following the “before” survey in the Birkenhead personalised marketing trial, households were divided into three “programme groups”:

- ▶ Group 1 (also known as “Programme Group A”): those who were interested in receiving information (excluding free PT tickets) and subsequently filled in an “information request form”
- ▶ Group 2 (“Programme Group B”): those households who were interested in receiving public transport, walking or cycling information (excluding free PT tickets) but did not complete an “information request form” (despite telephone reminders)
- ▶ Group 3: those who responded “no” to the two questions about receiving further information on public transport or walking and cycling.

Note, as indicated in Section 2.2.1, the basis for establishing these groups was completion of a 2-day travel diary and a positive response to one of the two information questions by at least one adult member of the household.

We analysed the first two groups (Programme Group A and Group B) to determine if there were any statistically significant differences between those who follow through on their expressed interest in receiving information and those who do not.

2.3.2. Findings

The number of household-related variables that can be compared between the two groups limited the analysis. The variables analysed included:

- ▶ Household size
- ▶ Number of adults in the household
- ▶ Presence of children in the house
- ▶ Number of children
- ▶ Number of vehicles per household
- ▶ Ratio of vehicles to adults
- ▶ Presence of a company vehicle in the household
- ▶ Distance to nearest bus stop and/or ferry terminal.

Only one of these variables was significant: the number of vehicles per household (refer Table 8). Group A households were more likely to have 1 or fewer vehicles (38% c/w 27%), $\chi^2(1)=6.745, p=0.009$.

Table 8 Number of vehicles per household by programme group

		Total N=458	Group A (receptive & requested info) N=230	Group B (receptive & did not request info) N=228
Number of vehicles usually at this address	1 or fewer	32.5%	38.3%	26.8%
	2 or more	67.5%	61.7%	73.2%
Total		100.0%	100.0%	100.0%

2.4. SELECTING THE TARGET AREA

2.4.1. International experience

Curtis and James (1998) and DFT (2002) conclude that it is important to focus on areas with good quality transport infrastructure for whichever mode the focus is on (i.e. PT, walking or cycling). However, it is imperative to realise that in most cases, areas are chosen where the PT service provision is of good or high quality (UITP/Socialdata 1998; Marinelli and Roth 2002; Kearns 1998). Furthermore, higher success rates apparently occur where the PT service provision is good *and* the current use is lower than could be expected for the area (DFT 2002). Marinelli and Roth (2002) also suggest that good support from stakeholders and PT operators helps.

Apart from the nature of PT services, very little else has been said about the characteristics of the physical locations where trials and demonstrations of personalised

marketing initiatives has occurred. Stopher (2003) compares the areas of Brisbane and Perth where IndiMark was implemented and found that the trial areas had similar characteristics: they were inner city, had lower car ownership than average, had smaller households with fewer children and more elderly people, and less driving to work than other parts of either city. However, it has also been suggested elsewhere that, in Perth, areas with low car ownership were actually more difficult to encourage to shift mode use (B. James, Transport Investment Consultant – City of Nottingham, personal communication, June 20, 2003).

With respect to car ownership, King and Salter (1997), trialling IndiMark in Hampshire County, also concluded that lower levels of car ownership might facilitate success. They found that where there were high levels of car ownership, it proved very difficult to motivate household members, with little experience of the bus, to give the bus a reasonable trial even when offered a free 1-month pass.

Stradling (2002), writing about reducing car dependence, would probably suggest that the above-mentioned Perth and Brisbane suburbs were less car-dependent than other suburbs. A car dependent *place* is one where it is more difficult to make alternatives to work, such as an area with high per capita vehicle ownership and use; low land use density; single-use land development patterns; large amounts of land for roads and parking; road designs favouring vehicle traffic; and reduced pedestrian environments. Interestingly, Stradling (2002) does not refer to the level of PT services in his classification of places as car dependent or not.

Geographic characteristics are discussed even less than the physical location characteristics. Marinelli and Roth (2002), in discussing the Grange Ward IndiMark® trial in Brisbane, thought that walking and cycling may have been unattractive due to geographical nature of area, which comprised undulating hills with many steep gradients.

2.4.2. Birkenhead location characteristics

In choosing the area for the demonstration programme, guidance was provided by the North Shore City Council, which had volunteered to be the sponsoring local authority for the demonstration. The Council considered Birkenhead to have a reasonable level of PT service – including buses and a ferry from Birkenhead Wharf to downtown Auckland – as well as to be suitable for walking and cycling, particularly given the extensive network of walking trails in the area (Booz Allen Hamilton, 2003). In addition, PT operators were also supportive of the concept.

In retrospect, however, Birkenhead may not have been the ideal choice for a demonstration programme. While the Council may perceive the level of PT service to be “reasonable”, it appears that at least some Birkenhead residents have a lower opinion of it:

“Just give me more buses, try improving the service itself, not the marketing aspect!!!”

“Stop marketing a useless system of transport”

“It proved to me that the bus service here is inadequate and antiquated”

“Bad service, uncomfortable, rude drivers”

“Buses too crowded and did not stay on timetable”.

The ferry operates half-hourly from Birkenhead Wharf during the morning and afternoon peak periods and hourly for the remainder of the time. It appears that *local* (within Birkenhead) bus services operate frequently (every 20 minutes during the weekdays), but many of the services into Auckland are less frequent – often only hourly, except perhaps in the peak periods when they operate every 20-30 minutes. Furthermore the services do not necessarily deliver people to their desired destination:

“The bus doesn’t go in the direction of where I work”

“If public transport was improved – later times, different routes – I might use it”

“No handy bus times to Takapuna.”

Although a “good” or “high” quality PT service is not well-defined in international experience, it appears that Birkenhead’s PT service may not meet this standard in the eyes of many residents. A successful personalised marketing initiative would involve a service that operates more frequently throughout the day, has routing to and from “common” destinations and has a good quality standard of delivery.

Topographically, Birkenhead suffers limitations as a good choice for the demonstration programme. Similar to the Brisbane trial, the area chosen is relatively hilly, which does not encourage walking and cycling. In addition, several people noted that cycling routes are too far away to be accessible and that “the roads are so dangerous here that nothing will get us to use one [a cycle]”. On the other hand, many positive comments were made about the provision of information on the walking tracks for recreational purposes. Again, this confirms international experience that if a desired outcome of a personalised marketing initiative is to encourage cycling (or walking), it is important to ensure that the infrastructure and environment for these activities is suitable and safe.

2.5. SUMMARY

The small number of households in the evaluative (“after”) survey pre-empts any analysis to identify significant characteristics of people who actually change their behaviour as the result of the personalised marketing initiative.

However, we did have a sufficient sample to analyse the receptiveness of *individuals* to a personalised marketing approach, wherein they were asked about their attitudes to alternative modes (specifically, public transport, walking and cycling) and whether or not they were interested in further information about these modes. We were also able to do a limited analysis of households to try to determine if there were any characteristics that differentiated those who followed through on their initial interest by filling in an “information request form.”

Personalised Marketing – Improving Evaluation

We found a number of characteristics to distinguish those individuals who were receptive to a personalised marketing initiative. For example, receptive individuals were more likely to:

- ▶ Have children
- ▶ Have 3 or more people living in the household
- ▶ Be aged between 15 to 44 years
- ▶ Be part time employees or students rather than retirees
- ▶ To have a lower vehicle / adult ratio (<0.5) in their household
- ▶ Not be “frequent” drivers (3-7 days per week)
- ▶ Already have some experience of PT use.

They were also somewhat more likely to be women. Receptive individuals also had significantly different responses to the attitude statements, being less likely to:

- ▶ State “nothing could induce me to walk any more than I do now”
- ▶ Agree or strongly agree with the statement “Even if PT was free I wouldn’t use it”
- ▶ Disagree / strongly disagree with the statements “I’d ride a bike at least once a week if I had a good bike and I knew there was a safe route” and “I like to walk at least once a day if possible.”

They were also more likely to demonstrate that they had been thinking about the pedestrian environment and to suggest ways to improve it (i.e. improve footpaths, crossings, lighting, etc.) so that they would be encouraged to walk more.

Only one characteristic carried through from the “receptiveness” to “taking action” (by filling in an “information request form”): households with 1 or fewer vehicles were more likely to follow through on their interest. No other household characteristic tested significant, including size, the presence of children or a company vehicle, and the distance to the bus stop or ferry terminal.

Although some of the characteristics show marked differences between receptive and non-receptive individuals, none of them are such that they could be classed as a “definitive” pre-selection criterion. For example, there is a 33% difference between those who were receptive and those who were non-receptive (64% c/w 31%) with respect to disagreeing / strongly disagreeing with the statement “Even if PT was free, I wouldn’t use it”. Even so, 17% (n=111) of the individuals who agreed or strongly agreed with this statement specifically requested information about the bus and ferry. Likewise, 40% (n=219) of those disagreeing /strongly disagreeing with the statement did *not* request any PT information. Classifying people in a personalised marketing initiative on the basis of their response to this statement could result in some interested individuals being left out, while other uninterested parties would be included.

The Birkenhead demonstration programme also highlights the need to ensure that a good quality PT service, along with good quality walking and cycling environments are in place prior to a serious attempt to market it to individuals. If people are to be encouraged

to use alternatives, then the alternatives must be adequate to meet their needs. In terms of geography, hilly terrains are unlikely to generate good walking and/or cycling results.

3. EVALUATION PITFALLS

As noted earlier, a lot of the international literature on personalised marketing initiatives is focused on measuring the behavioural change associated with the programme. Unfortunately, the evaluation has not always been as rigorously executed as is required to be able to confidently measure the behavioural change. This is not an uncommon fault in evaluation – standard texts such as Rossi and Freeman (1993) observe:

“We cannot overemphasize the technical and managerial difficulties involved in undertaking impact evaluations.... The problem of establishing a programme’s impact is identical to the problem of establishing that the programme is a cause of some specified effect.” (p. 218)

They note that establishing such causal relationships is confounded by biases in the selection of participants, particularly where a programme is voluntary (as in this case), and by changes in external factors (i.e. change in season or weather pattern between the “before” to “after” surveys; introduction of new PT services, road works, special events such as “cycle to work day”).

Completing the “perfect” evaluation of interventions such as personalised marketing may be nearly impossible, given the constraints of budget, limited respondent patience, etc. However, there are some basic rules to undertaking a “good enough” evaluation, including choosing the appropriate sample size to take account of known variability, reporting reasonable statistical confidence intervals or statistical tests about the (possible) behavioural change, taking care with the logic of experimental or quasi-experimental design, and acknowledging underlying trends in mode use. Unfortunately, such basic rules appear to have been ignored in some evaluations of personalised marketing.

Richardson (2003) provides an extensive explanation concerning variability and sample size. Our work complements his work by focusing on:

- ▶ new information about variability and sample size derived from a major New Zealand survey
- ▶ the impact of pre-existing differences on “before” and “after” results in such surveys.

3.1. CHOOSING APPROPRIATE SAMPLE SIZE

Near the beginning of the Birkenhead demonstration programme, we were asked to provide analytical inputs into the sample and evaluation methodology. The key driver of sample size estimates is the measurement of “variability” or consistency of the behaviour under question (i.e. number of trips taken or distance travelled in a day). In short, relatively high variability results in large samples being necessary, and low variability allows smaller sample sizes.

In order to approximate the sample size required for the demonstration programme, we used a major nationwide travel behaviour survey (the 1997/98 New Zealand Travel

Survey conducted by the Land Transport Safety Authority), which has a distinctly high response rate (75% of eligible dwellings yielded full response from all household members). Fortunately, this survey records travel behaviour on two days rather than the more common one-day travel diary; hence estimates of variability between days are possible. The two days recorded are consecutive, meaning that our estimates only approximate the differences to be expected from travel diaries used for personalised marketing (which are usually several months apart, and should be matched for the same day of week). We restricted our analysis to a sub-sample from Auckland (as most relevant to the planned Birkenhead personalised marketing trial) and to weekdays only (because the main impact analyses for the demonstration programme were to be focused on weekdays not weekends, and also because mixing together weekdays with weekends would increase variability). Calculations are presented for two fundamental methodological options:

- ▶ **Independent groups.** This concerns two separate groups of people such as marketing group versus control group, or general survey of area before personalised marketing versus marketing group (sometimes referred to as a "repeated cross-sectional" approach)
- ▶ **Panel survey.** This involves the same people measured at least twice (e.g., before personalised marketing versus after personalised marketing).

Our measure of variability is the coefficient of variation of the number of relevant trips, that is, the standard deviation divided by the mean⁵. We reduced the samples to respondents aged 15 years and over only (consistent with practical data collection) and to households where all eligible respondents completed all survey forms (for clarity with respect to our results concerning appropriate numbers of households). Together with the earlier restrictions (Auckland only, weekdays only), this resulted in a sample size of 832 adults (15 years +) in 417 Auckland households.

As expected, Table 9 shows that the day-to-day variability for panel surveys is consistently lower than for independent groups. Furthermore, variability for public transport travel is distinctly greater than for driving. The practical implication is that markedly greater sample sizes would be required to detect a change of the same size for public transport travel compared with driving. Although the focus of this trial is on driving and public transport use, we also completed some illustrative variability calculations for other modes. For the panel survey approach (relevant to the Birkenhead trial), these results were car passenger 270%, walk 185%, and bicycle 1273%. Cycling behaviour appears *extremely* variable and hence it may be particularly difficult to reliably assess changes in cycling (unless the changes of interest are exceptionally large).

⁵ For the panel survey, the relevant measure is the standard deviation of the **difference** between the two days recorded (day1–day2) divided by the mean for the two days.

Table 9 Variability between weekdays

		Person	Household
Independent groups	Drive	102%	89%
	Bus or train	449%	345%
Panel survey	Drive	82%	56%
	Bus or train	277%	196%

Assuming that the difference to be detected for car driving is 10% and 20% for PT use, which seems reasonable given results reported from South Perth (Brög and John, 2001), Table 10 shows the sample sizes required to measure change in travel patterns on weekdays only. Note that these sample sizes for the panel survey concern people and households from which data has been successfully collected both “before” and “after”, not merely people/households contacted at the start (many of whom might refuse to respond).

Table 10 Approximate sample sizes required to reliably detect 10% change in car driving trips or 20% change in public transport trips on weekdays (1-day trip diary)

		Person	Household
Independent groups	Drive	2252	1705
	Bus & train	8000+	6449
Panel survey (without control group)	Drive	727	345
	Bus & train	2071	1038

The first important conclusion is that sample sizes required to reliably test for changes may be much greater than those implied by widely-cited local studies on this topic such as the sample sizes of around 200 each for the marketing and control groups in the original South Perth trial using independent groups (Brög and John, 2001). Sample size requirements may be somewhat lower if units of measurement differ from the trip legs which are the fundamental unit of the database we used (a journey to work which includes a brief stop to drop off a passenger counts as two trip legs). Note that Richardson (2003) showed that sample size requirements were consistently higher for vehicle kilometres than vehicle trips (and our preliminary calculations confirmed that this inconvenient pattern also held in New Zealand, even when extreme values from very long trips were excluded).

Guidelines for appropriate sample size should have been readily available from previous personalised marketing studies. Unfortunately, however, the sample size adequacy in the Australian studies cited was impossible to assess at the time we were planning the Birkenhead study because of the repeated absence of statistical significance tests in the published reports (contrary to professional practice established in codes of conduct for survey research both locally and internationally). Nor did direct contact with Perth in early 2002 deliver us sound examples of statistical testing. Doubt over whether the widely cited South Perth results provide any statistically robust evidence of IndiMark impact was particularly disturbing given the high benefit-cost ratios claimed. For example, Brög and John (2001) asserted a ratio of 13:1 for the pilot IndiMark programme, and 30:1 or higher for the programme extension to half of the Perth Metropolitan Region.

Second, required sample sizes may differ substantially depending on the transport mode of interest (as illustrated for bus and train above).

Third, although it is no surprise that measurements averaging across all people in a household require lower sample sizes, the results here quantify this (which is important given the greater effort required to get complete responses from all eligible respondents in a household).

Fourth, these results highlight the importance of considering alternative measurements to the common one-day travel diary in order to reduce variability (and hence sample size requirements), despite the risk of lower response rate from greater demands of respondents. For example, one may wish to consider travel diaries for 2–7 days, or “odometer surveys”, recording household vehicle kilometres travelled before and after a personalised marketing programme.

In parallel, as we were planning the Auckland study, Richardson (2003) was completing similar calculations to contribute to a personalised marketing trial in Melbourne. Similarities and differences in both results and method are instructive (Richardson's paper provides details of formulas used for sample size calculations.) Firstly, both sets of results highlight the possibility of large sample sizes much greater than 200 being required with the conventional one-day travel diary. Secondly, the fundamental estimates of variability are broadly similar despite very different data sources. For example, our 89% and 56% (Table 9) are remarkably close to Richardson's 85% and 60% (Table 3, daily trips per household). Estimates of sample size requirements differ more because of different assumptions in calculations discussed below, e.g., our 1705 (Table 10) compares with Richardson's 762 (Table 8, repeated cross-sectional survey, household, trips).

The broad similarity in the estimates of variability is reassuring. They occurred despite several major underlying differences including:

- ▶ Our estimates are directly from 417 Auckland households, whereas Richardson's source is 146 German households adjusted substantially to fit Melbourne (e.g., to take account of Melbourne people making around 42% more car trips per week).
- ▶ Our data concerns the differences between two consecutive weekdays, whereas Richardson's concerns all days including weekends. A clear strength of Richardson's data is his variability estimates from matched the days of the week (i.e., looking at differences between Tuesdays etc) and also from data collected over complete weeks rather than single days.

Our sample size estimates differ somewhat more from Richardson than our variability estimates, particularly with respect to larger sample sizes, because we chose not to use the Finite Population Correction Factor (FPCF). The use of the FPCF ensured that Richardson's sample size requirements could never exceed 1500. Using the FPCF is conventional for estimating margins of error where the total population of households being considered is as low as 1500. We chose not to use it, not just because the population in Birkenhead was somewhat larger, but because we believe that the relevant population of interest is the many broadly comparable suburbs in Auckland urban area rather than Birkenhead alone. Such a large target population makes the FPCF irrelevant in practice (the restriction to Birkenhead for the demonstration programme is purely a

matter of pragmatic convenience, as was perhaps the area chosen in Melbourne). In short, it was not of interest to test whether we could statistically prove that a difference had occurred in a relatively small area – if one collects data from a sufficiently large proportion of such a limited population, even totally trivial differences will be statistically significant⁶. This argument parallels the difference noted by the statistician Deming (1975) between "enumerative" and "analytic" studies. In addition, we do not wish to publish sample size requirements that might mislead the later researchers working with larger populations than ours into using lower sample sizes than desirable because of a largely hidden background factor.

As a result of our calculations of variability and sample size, the Birkenhead demonstration programme used a panel survey approach with a matched two-day travel diary in an attempt to reduce the variability (that is the same two days were recorded by respondents in the before and after surveys). However, two factors undermined the beneficial effects of this: (1) the increased size of the survey meant that the number of households in the initial sample had to be reduced to keep within budget; (2) the more onerous task of completing 2-day travel diaries for an entire household contributed to the response rate being much lower than what was originally expected (632 households responded out of 2100 approached in the "before" travel diary, with 595 completing the 2-day travel diaries, while only 139 households completed the "after" travel diary.) Thus, from the outset, we did not expect to bridge the large gap between the ideal sample size for significance testing and what was affordable in the Birkenhead study trial. The high levels of variability and their effect on sample size requirements mean that using traditional travel diary methods to quantitatively demonstrate the impact of personalised marketing will often not be cost-effective in small pilot projects.

3.1.1. Appropriate reporting of statistical significance

Sufficient sample size is vital to being able to report statistically robust results. In addition, there are standard professional practices which should be adhered to in the reporting of statistics. Many of the evaluations of IndiMark trials in Australia and New Zealand have not demonstrated these common practices.

Before going on, we want to make it clear that we accept that the broad pattern of results strongly suggest that individualised marketing is a very promising approach to reducing car driver trips and definitely worthy of further effort. The evidence from the electronic ticketing system confirms that public transport use in South Perth increased with individualised marketing (Roth et al., 2003). However, given the general propensity to "prove" the impact of IndiMark with sample sizes too small to be statistically robust, we would urge caution in making claims about the magnitude of the change associated with such a programme. Poor reporting practices mean that there has often been no indication of the degree of uncertainty surrounding the published results.

Recent reports, presentations and papers (e.g. Socialdata 2003a and Roth et al., 2003) have rebutted some criticisms of the validity of individualised marketing results in Perth such as those by Stopher and Bullock (2003). However, the fact remains that the issue

⁶ In the extreme, imagine that results were obtained from nearly all 1500 households in the region for a personalised marketing demonstration, and that they showed a tiny difference such as a reduction in car trips of 1%. This difference would be statistically significant because it is based on data from nearly the entire population. But it would be far from encouraging as practical evidence to extend the trial to the larger surrounding urban area.

of appropriate sample size and appropriate reporting of statistical results and tests has been poorly handled, repeatedly in the case of Perth, as well as in published reports of Brisbane, Cambridge and Auckland studies (James, 1998; Brög & John, 2001; Socialdata 2003b; Booz Allen Hamilton, 2003). There are several practical lessons from this for future evaluations if they wish to be convincing rather than attract the critical comment that Perth has.

The repeated failure to clearly report base sample sizes, including the split between different groups within the sample, and the lack of indications of statistical significance is particularly disturbing given that, *internationally*, this is not acceptable professional practice with reporting on surveys. For example, Rule 25 of the ICC/ESOMAR International Code of Marketing and Social Research Practice (ESOMAR 2003) states quite clearly:

"The client is entitled to...

- Bases of percentages (both weighted and unweighted)
- General indicators of the probable statistical margins of error to be attached to the main findings, and of the levels of statistical significance of differences between key figure"

The ESOMAR code is the foundation for the Codes of Practice of both the Australian and New Zealand Market Research Societies. Putting aside papers for publication (which, for the sake of professionalism, should adhere to the same rules), even in reports written specifically for clients (e.g., Socialdata, 2003b; Booz Allen Hamilton, 2003), there has been a failure to adhere to these basic practices.

The practical implications of failure to adhere to such professional standards typically include waste of time, effort and money in the long run. For example, had the earlier Perth reporting and analysis followed professional conventions, the general conclusion that larger sample sizes than expected were needed for solid survey-based quantitative impact evaluation of such initiatives could have been reached in 1998 or 1999 rather than 2003. Given the low cost-effectiveness of such extensive survey requirements, this may have resulted in the adoption of other evaluation techniques for more recent individualised marketing projects (e.g. Auckland and Melbourne), thus saving time, effort and money.

There have been some recent attempts to produce significance tests for South Perth survey results (namely, Goulias et al., 2002; Socialdata 2003a; Roth et al., 2003). However, these serve to raise further concerns about the research process used in South Perth and certainly do not demonstrate good procedures suitable for use by others. Contrary to the conclusions of Goulias et al. (2002), their tests do not convincingly demonstrate impact of personalised marketing. Instead, they largely show unsurprising pre-existing differences similar to those discussed in section 3.3 below (e.g., it is to be expected that the group classified at the very start as regular users of public transport will use public transport significantly more and drive cars significantly less than other respondents).

More recently, Socialdata Australia (2003a) has produced a Technical Appendix which includes detailed calculations for significance tests on mode choice for both persons and trips. Rather than allay our concerns, these calculations only serve to amplify them,

because every calculation is based on whole percentage changes – i.e. taking no account of rounding. For example, consider the test for bicycle mode choice based on persons and trips: the test calculations use a change of exactly 1% as the difference of interest (2% before compared with 3% afterwards). Likewise, the change in public transport trips uses a change of exactly 1%. If the underlying results for bicycle mode are mode shares of say 2.4% and 2.6%, then the test result would have been close to 0 and very clearly not statistically significant (rather than the highly significant result reported). We argue that the real change in bicycle mode share or public transport share was highly unlikely to be exactly 1.00% or that the change in other modes were consistently whole percentage changes – statistically, this would be extremely improbable..

In summary, we urge greater openness with respect to reporting weaknesses rather than inappropriate claims of rigour such as those made by Brög and John (2001, p.5). It only takes a few words to mention that results are indicative only rather than statistically robust, or to show awareness of stronger methods. Such realism helps cumulative improvement of quality with such research in the future, in addition to allaying concerns that those designing and using the study are dangerously ignorant of the weaknesses. Only many years after the 1997/98 surveys for Perth did we see significance tests for the survey results, and only later were there frank admissions that the cross-sectional design (different respondents in the "before" and "after" samples) was a weakness and that panel survey designs using the same respondents "before" and "after" are the preferred method (Goulias et al., 2002; Socialdata Australia 2003a). Had this type of information been available at the outset, much angst would have been foregone.

3.2. CONTROL GROUPS

To establish that the personalised marketing trial has had an impact, one needs to compare travel behaviour of those exposed to marketing with some suitable comparison group. There are a variety of possibilities, and few easy answers in practice.

Those planning such evaluations need to ensure that the team includes someone confident to assess trade-offs between experimental design options such as:

1. Control groups (a no-treatment group for comparison with those exposed to marketing)
 - ▶ Broadly, this was the approach used in South Perth
2. Panel survey (using the same respondents before and after personalised marketing, "using respondents as their own controls")
 - ▶ A major advantage compared with using control groups of different people is substantial reduction in required sample size
3. Panel survey + control group (the logical ideal which combines the strengths of both of the previous options)
 - ▶ Broadly, this was the approach used recently in Brisbane (Marinelli and Roth, 2002)
4. Other alternatives
 - ▶ For example, in Birkenhead a panel survey approach was used but cheaper methods were used to avoid the extra data collection costs associated with

having a control group as well. The cheaper methods attempted to rule out major factors likely to change travel behaviour by gathering background data (traffic counts, public transport patronage data, walking and cycling counts, and even weather records).

3.3. ESTABLISHING “BEFORE” AND “AFTER” TRAVEL PATTERNS

Recently, Goulias et al. (2002) published a paper, which claims to demonstrate that the IndiMark® programme was responsible for the distinctive mode share behaviour of the so-called “Interested” group (the group targeted for receiving personalised marketing) in the 1997 Perth trial. They also note that the method of data collection (constrained by funding) meant that there was no “microstate analysis to study individual-by-individual net change of behaviour”. Unfortunately, this inability may have generated quite misleading conclusions.

In section 2.2.2, we discussed the differences in PT use and car driving patterns between receptive and non-receptive individuals. These differences were apparent from the outset of the programme – that is *before any personalised marketing approach had been applied*. Once the hierarchical classification of “Interested”, “Regular” and “Not interested” was applied, we found that, of the 235 individuals who used PT one or more times per week, 195 were in the “programme group” of households, with the remainder in the “regular” and “not interested” groups. These 195 individuals comprised 20% of the target population in the programme households.

The presence of infrequent and frequent PT users in the programme groups is not, in itself, surprising, given the findings of various studies that people who have had some interest in, and exposure to, alternative modes are more likely to be approachable for further change (see for example, DFT 2002). What is of interest, however, is the extent of the difference in their mode use.

Both infrequent and regular PT users were far more likely to be car passengers one or more times per week than non-PT users (71% c/w 43% of non-PT users, $\chi^2(2)=54.3$, $p<0.001$) and to drive a car less frequently (86% of non-PT users drove 3-7 days per week c/w 31% of PT users, $\chi^2(2)=218.4$, $p<0.001$). Such dramatic differences in PT, car driver, and car passenger travel patterns by 20% of the adult population in the programme households will undoubtedly significantly influence the overall travel patterns of the household.

In fact, this is exactly what the Birkenhead data suggests when the total number of journeys made by households, as well as the number of journeys⁷ incorporating the use of car (both as driver and as passenger) and public transport is generated. However, due to respondent bias (the households are self-selected to participate and, as a result, 72% are in the “receptive” group and are the key drivers of the mean results for the total sample), we must stress that these results cannot be treated as statistically robust.

⁷ In the self-completion travel diary, respondents were asked to “record any journey more than 100 metres long on public roads or footpaths – even short journeys like walking to lunch and back may be important. A journey ends when you GO HOME or STOP for 1 hour or more.” The mode use within each journey was recorded in the total number of minutes for each mode used, not by the number of times a mode was used in a journey (i.e. walking – bus – walking would be recorded as walking=x minutes; bus=y minutes).

Nevertheless, we provide the results in Table 11 to provide an indication of what the travel patterns of the three household groupings could look like *before any personalised marketing occurs*.

Table 11 Proportion of total journeys involving use of specified mode (driver, passenger, PT) by "interested", "regular PT user" and "not interested" households

Mode	Share of total journeys (%) ^a			
	Interested	Regular PT users	Not interested	Total
	<i>N=445 households & 4663 journeys</i>	<i>N=38 households & 340 journeys</i>	<i>N=111 households & 851 journeys</i>	<i>N=594 households & 5854 journeys</i>
Car driver	74.0	51.4	82.8	74.9
Car passenger	12.6	15.0	10.2	12.1
Public transport (bus & ferry)	8.0	34.1	1.2	7.6

^a Totals do not add to 100% because not all modes used are included in the table.

These observations stress the importance of accurate data collection from the same group of people *before* and *after* the programme (or at least from reasonably comparable groups of people who have been accurately classified into comparable groups) in order to evaluate the impact of a personalised marketing programme.

Hence, contrary to conclusions of Goulias et al. (2002), we do not see their re-analysis as providing any statistically reliable evidence that IndiMark resulted in changed transport behaviour in South Perth. Instead, their results appear easily explained in terms of possible pre-existing differences between groups.

4. CONCLUSIONS

Our analysis and experience from the Birkenhead personalised marketing demonstration programme, leads us to provide the following guidance to local authorities and others who may wish to apply a personalised marketing programme in a given area and to measure the impact of such an intervention:

Pre-selection criteria

- ▶ We found that the people who were receptive to a personalised marketing approach were different to those who were not (i.e. they had more positive attitudes to alternative modes, tended to have had some experience of PT, to come from larger households, not be retired, and have a lower adult:vehicle ratio). Although all of these differences are statistically significant, no one descriptor should be seen as the “definitive” characteristic to rely on when selecting individuals or households for participation in a programme.
- ▶ Good quality infrastructure and services for the alternative modes (PT, walking and/or cycling) that will be the focus of the personalised marketing programme is important. A “reasonable” quality PT service from the perspective of a local council may not be “good enough” for the area where marketing will occur. Frequency, routing and service quality (friendly drivers, timeliness) are all important features. If roads are too busy or dangerous for cycling or the cycle paths are located too far away from the target area, then there is likely to be a negative response to encouraging people to use them.

- ▶ Areas that are topographically “challenged” (i.e. hilly or with very long distances between desired destinations) may not be the most suitable for personalised marketing programmes.

Evaluation pitfalls

- ▶ There is a great deal of variability in day-to-day use of modes by individuals and households. This means that the sample sizes required for travel behaviour surveys to firmly establish any causal effect of a personalised marketing programme are substantially larger than has generally been the case with trials to date in Australia and in New Zealand. This suggests that other evaluation tools to measure such travel behaviour change (such as odometer surveys or collection of electronic passenger boarding data) may be more practical and cost-effective.
- ▶ Along with being large enough to detect significance, care must be taken when evaluating results to ensure that the group composition is the same for the before and after surveys, as we have demonstrated that the “interested” group is highly likely to have different mode shares *at the outset of the programme*. This is a further confounding factor (to others such as the weather, change in seasons, road works, changes in PT service, underlying trends in mode use, etc) that may affect the interpretation of results.
- ▶ Great care is required in selecting an appropriate control group for comparisons.
- ▶ Contrary to some recent examples of reporting and analysis, it is important to clearly show the base number of respondents in major groups and also to report margins of error or statistical significance tests for important changes found.
- ▶ It is good practice to briefly mention if results are intended only to be illustrative or indicative rather than strong claims implying statistical significance, and also to briefly mention important design weaknesses or suggestions for improvement in future. Such habits are important both to allay suspicions that authors of reports are unaware of weaknesses, and also to contribute to cumulative improvement in research quality.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial assistance provided by the Foundation for Research Science and Technology without which this project, as part of a larger research programme, could not have been undertaken. We also acknowledge and thank the many “end user” collaborators who helped us to focus our research programme into the areas of greatest interest. Without people and organisations to support our efforts and to use our results, there would be no reason to begin any research programme.

REFERENCES

Booz Allen Hamilton and Pinnacle Research. June 2003. Home Travel Plans.
Unpublished report prepared for the Energy Efficiency and Conservation
Authority, Wellington, New Zealand.

- Booz Allen Hamilton. 2003. Personalised marketing demonstration study for Birkenhead Auckland. Transfund New Zealand Research Report 246. Wellington, New Zealand.
- Brög, W, & John, G. 2001. 'Personalised Marketing – the Perth Success Story' In Proceedings: *Marketing Public Transport Conference*, Auckland, New Zealand, 3 August 2001. www.dpi.wa.gov.au/travelsmart/techpub.html
- Campbell, DT, & Stanley, JC. 1966. *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- Curtis, C & Headicar, P. 1997. Targeting travel awareness campaigns: which individuals are more likely to switch from car to other transport for the journey to work? *Transport Policy* 4(1): 57-65.
- Deming, W.E. 1975. On probability as a basis for action. *The American Statistician*, 29, 146-152.
- Department for Transport (UK). 2002. A Review of the Effectiveness of Personalised Journey Planning Techniques. Accessed on 27 February 2003 at: <http://www.local-transport.dft.gov.uk/travelplans/pjourney>
- ESOMAR. 2003. ICC/ESOMAR International Code of Marketing and Social Research Practice. www.esomar.org (Codes & guidelines; accessed 29 Oct 2003).
- Goulias, K.G., Brög, W.W, James, B. and Graham, C. 2002. Travel behaviour analysis of South Perth individualised marketing intervention. *Transportation Research Record* 1807, pp. 77-86.
- King, P and Salter, S 1997. Switching to Public Transport: Results of a UITP Trial. PTRC Seminar G.
- Marinelli, P. & Roth, M. 2002. TravelSmart suburbs Brisbane – a successful pilot of a voluntary travel behaviour change technique. In Proceedings: 25th ATRF, 2-4 October 2002, Canberra, Australia.
- O'Fallon, C; Sullivan, C and Hensher, D. in press. Constraints affecting choices by morning car commuters. *Transport Policy*. Further details available from www.pinnaclereasearch.co.nz
- Roth, M. Ker, I., James, B., Brög, W, Ashton-Graham, C., Ryle, J, Goulias, K., & Emmerson, R. A dialogue on individualised marketing: Addressing misperceptions. Paper presented to the 26th Australasian Transportation Research Forum, 1-3 October 2003, Wellington New Zealand.
- Richardson, A.J. (2003). "Temporal Variability of Car Usage as an Input to the Design of Before & After Surveys", Paper Presented at the 82nd Annual Meeting of the Transportation Research Board, Washington, D.C.

Socialdata Australia. 2003a. Technical Appendix (Travel Behaviour Change Program for the City of South Perth under the TravelSmart Program 2001-2005). www.dpi.wa.gov.au/travelmart/techpub.html (accessed 29 Oct 2003).

Socialdata Australia. 2003b. Final Report – Town of Cambridge. Travel behaviour change programme for parts of the Perth Metropolitan Area under the Travel Smart Program 2001-2005. www.dpi.wa.gov.au/travelmart/ (accessed 4 Nov 2003).

Stopher, P. & Bullock, P. 2003. Travel behaviour modification: a critical appraisal. Paper presented to the 26th Australasian Transportation Research Forum, 1-3 October 2003, Wellington New Zealand.

TAPESTRY (Transport & Travel Research Ltd et al.). November 2001. Deliverable 2 – State-of-the-art review. Project funded by the European Commission, Transport RTD Programme, 5th Framework Programme. Accessed on 7 May 2003 at: http://www.eu-tapestry.org/p_dwl/tap_d2.pdf

UITP / Socialdata. 1998. Switching to Public Transport. Europe: UITP.