

A decade of evaluating voluntary travel behaviour change: why don't we know what works?

Charles Sullivan

Australasian Evaluation Society Conference, Wellington, September 2010



Kinnect
group

Capital Research • Judy Oakden Consultancy
Research Evaluation Consultancy • The Knowledge Institute
Julian King and Associates Limited

Background



Kinnect
group

Large investment in roading, but outcomes?



☺✓ Walking school buses

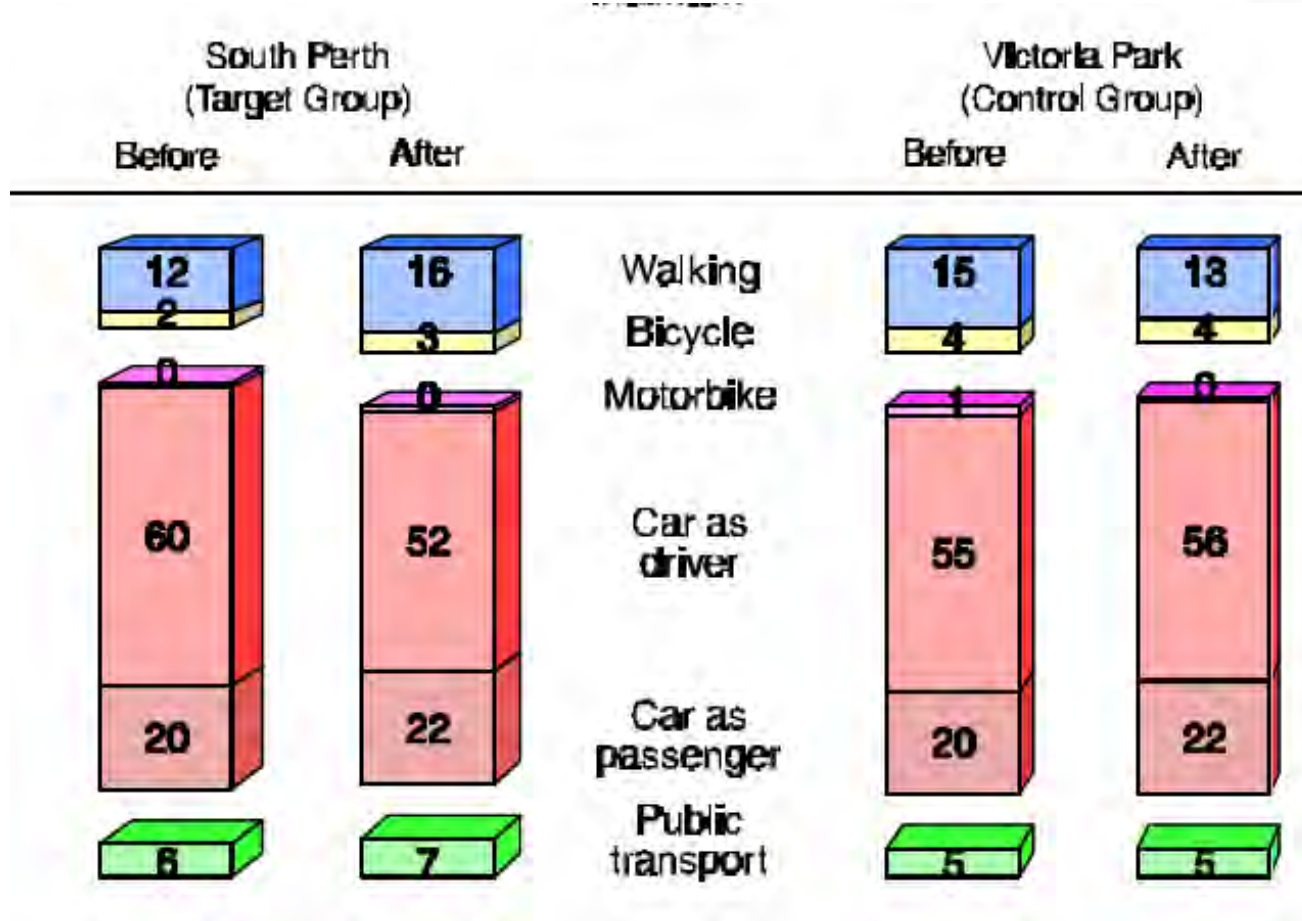


😊? Workplace travel plans

- Systematic efforts tailored to a workplace
 - E.g. cash instead of parking, carpooling support (coordination, priority car parks, guaranteed ride home), cheaper public transport, showers & safe storage for cyclists...

“Workplace travel plans typically reduce commuter car driving by between 10% and 30%, though the best ones achieve significantly more than that. Typical cost to the local authority is £2-£4 per head.” Cairns et al., 2004, p. 365; *Smarter Choices* report to Department for Transport (UK)

☺? Individualised marketing—the Perth success story (1)



SocialData (2001)

😊? Individualised marketing—the Perth success story (2)

- Net present value of a proposal to deliver individualised marketing in the broader Perth region ... over AUS\$1 billion, with a benefit-cost ratio of at least 30:1 (Brög & John, 2001)

... and lots more evaluations of household individualised marketing in Australia alone

There is little further need to undertake major evaluations of household projects, as the Australian and international data is in broad agreement, and there is little need to demonstrate the effectiveness of methods used”.

(Australian Greenhouse Office, 2005, p. 5)

☹ But Nov 2009 Special issue of major journal



“Evaluation of programmes promoting voluntary change in travel behaviour”



☹️ Title: *Do we know whether personal travel planning really works?*

- “Unfortunately, despite a decade of experience with PTP we still cannot be sure whether it works.” (Bonsall, 2009, p.312) [UK]

☹️ Title: *Evaluating personal travel planning: If it is prohibitively expensive to get a robust answer then what should we do?*

- Cohen (2009) [UK]

... and more generally (incl. travel plans)



The overall conclusion from the above is that the collection of information on the emergence of environmentally sustainable travel and activity

patterns is **fraught with difficulties**. It seems that no single technique can be relied on to give a complete and unbiased picture ... The obvious implication is that analysts should proceed with care and, where possible, try to use data from a variety of sources (preferably ones which are unlikely to suffer from the same biases!).

(Bonsall, 2008, p. 21)

Reflections: Why?



Kinnect
group

Questions

- Why have we done so much monitoring/ evaluation and learned so little about VTBC impact in the last decade?
- Why do unsatisfactory methods get repeated again and again and again and again and again and again and again and again...?
- When shouldn't an impact evaluation be done?
- Why did it take so long for the UK to learn what we knew by 2002/03?

Evaluation not independent

- UK (mainly 2005 onwards)
16 of 23 case studies in
8 UK areas had same
evaluator, SocialData
 - “evaluation of most of the well known PTP schemes was conducted by a company with a commercial interest in demonstrating the success of PTP [SocialData] ... it would have been better for all concerned had the evaluations been conducted by an independent body” (Bonsall, 2009, p.312)
- Melbourne 2003
 - Evaluation of several major new PTP projects independent of implementation (cf. Perth where SocialData did both)



Statistical fundamentals ignored

- UK 2009: “The size of samples in the project surveys has been insufficient”
 - (Chatterjee, 2009; p. 301; reviewing 23 case studies in 8 UK areas--mainly 2005 onwards)
- Wellington 2003 (ATRF conference): Surprisingly large sample sizes needed to detect likely differences in typical PTP surveys
 - Used to plan subsequent Australian sample sizes



Comms failure 1: statistical significance

- In 2002 we could not get significance test results for 'Perth success story' ('after' survey in 1998)
 - \Rightarrow hard to plan sample size for NZ pilot of PTP...
- NZ 'demonstration project' in Birkenhead reported 4.5% decrease in car driver trips "similar in to those found in other personalised marketing experiments"
 - significance test and sample size problems not mentioned...
- RFP for second NZ PTP study came out with inadequate sample size

...

- Several UK local authorities used same core RFP for PTP projects
 - Around 10% of £150,000 budget for monitoring/impact
 - £15,000 clearly inadequate to deliver what they expected
 - Cohen (2009) reports attempting to explain the inadequacy to clients “with only limited success”
- UK review: “statistical significance of changes in travel behaviour ... was not always reported ... where it was reported, there are some concerns about the basis for the tests used” (Chatterjee, 2009, p. 299)

Whose professional standards apply?

- “The client is entitled to ...General indicators of the probable statistical margins of error... and of the levels of statistical significance of differences between key figures” (ICC/ESOMAR 2003)
- But do other professionals comply when reporting surveys? Should they?
 - Transport consultants?
 - Marketers?
 - Evaluators?

Comms failure 2: success trumpeted more



success!

But

*No impact
Rebuttal*

Comms failure 3: reports hidden from Google

- Western Australia very good at putting technical reports on the web
 - ≈ 30 on PTP; 3 on design of workplace travel surveys
- Victoria has ≈ 30 reports on travel behaviour change, incl. 10 on evaluation
- NZ transport sector often doesn't do this
 - Inconsistent with NZ reliance on literature review to guide plans and technical decisions
 - Hinders cumulative growth of knowledge
 - Fosters repetition of mistakes



Uncritical repetition of method

A supine acceptance of 'before and after' self-completion travel diaries as the method of choice is holding us all...back

Cohen, 2009, p. 346

- 'Sharp' debate/ dialogue session on PTP evaluation at 2003 ATRF conference
- Australian PTP evaluation diversified after 2003
 - Bus/train boarding trends, odometer surveys, personal GPS



Reviews not critical enough?

“There is little further need to undertake major evaluations of household projects, as the Australian and international data is in broad agreement, and there is little need to demonstrate the effectiveness of methods used”

Australian Greenhouse Office,
2005, p. 5

- In contrast, UK 2009
 - “although the prominently published results show a degree of consistency, serious questions remain about the reliability of the methods used to produce these results and about the possibility of systematic bias. It is argued that the possibility of systematic bias undermines the weight that would otherwise be given to the consistency in published results” (Bonsall, 2009, p. 306)

Conclusions

Don'ts and do's



Kinnect
group

Don't:

- Ignore sound advice from a 'different' profession
- Ignore statistical criticisms because conclusions are inconvenient
- Ignore sound advice or precedents from smaller countries
- Repeat previously used methods without critical appraisal of weaknesses
- Insist on impact evaluations with demonstrably insufficient budget
- Commission ineffective monitoring/ evaluation largely to make it easier to get agreement for core project

Do:

- Publish technical evaluation design documents on the web
- Report limitations (e.g. non-significant differences, insufficient sample size, bias)
- Put 'unsuccessful' impact evaluations on the web
- Be prepared to pay extra to have evaluation independent of implementation
- Put in the extra effort at conferences for debates, dialogue sessions, or other direct opposition of views

For further information contact:

Charles Sullivan

Email: charles@kinnect.co.nz

www.kinnect.co.nz

Bibliography

- AEA Evaluation Policy Task Force. (2009). *An evaluation roadmap for a more effective government*. Washington, DC: American Evaluation Association. Retrieved from www.eval.org/aea09.eptf.eval.roadmapF.pdf
- Australian Greenhouse Office. (2005). *Evaluation of TravelSmart projects in the ACT, South Australia, Queensland, Victoria and Western Australia: 200 1–2005*. Canberra: Australian Greenhouse Office.
- Bonsall, P. (2009). Do we know whether personal travel planning really works? *Transport Policy*, 16(6), 306-314. doi:10.1016/j.tranpol.2009.10.002
- Booz Allen Hamilton. (2003). *Personalised marketing demonstration study for Birkenhead Auckland* (No. 246). Wellington: NZ Transport Agency.
- Brög, W., Erl, E., Ker, I., Ryle, J., & Wall, R. (2009). Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*, 16(6), 281-292. doi:10.1016/j.tranpol.2009.10.003
- Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A., & Goodwin, P. (2004). *Smarter Choices – Changing the way we travel*. London: Department for Transport,
- Chatterjee, K. (2009). A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*, 16(6), 293-305. doi:10.1016/j.tranpol.2009.10.004
- Cohen, T. (2009). Evaluating personal travel planning: If it is prohibitively expensive to get a robust answer then what should we do? *Transport Policy*, 16(6), 344-347. doi:10.1016/j.tranpol.2009.10.005
- Goulias, K., Brög, W., James, B., & Graham, C. (2002). Travel behaviour analysis of South Perth individualised marketing intervention. *Transportation Research Record*, 1807, 77-86.
- O'Fallon, C., & Sullivan, C. (2004). Personalised marketing—improving evaluation. *Transport Engineering in Australia*, 9(2), 85–102.
- O'Fallon, C., Sullivan, C., & Cottam, P. (2002). Walking School bus networks: A 'flaxroots' approach to cleaner air. Presented at the 16th International Clean Air & Environment Conference,, Christchurch. Retrieved from www.pinnacresearch.co.nz/research/WSB/O'Fallon2002WSB.pdf
- Richardson, A. J., & Harbutt,, P. L. (2003). Design issues for before and after surveys of travel behaviour change. Presented at the 26th Australasian Transportation Research Forum, Wellington.
- Richardson, A. J., Roddis, S., Arblaster, D., Attwood, D., & Newman, J. (2005). The role of trend analysis in the evaluation of a TravelSmart program. Presented at the 28th Australasian Transport Research Forum, Sydney.
- Roth, M., Ker, I., James, B., Brög, W., Ashton-Graham, C., Ryle, J., Goulias, K., et al. (2003). A dialogue on individualised marketing: Addressing misperceptions. Presented at the 26th Australasian Transportation Research Forum, Wellington.
- Stern, E. (2009). Evaluation policy in the European Union and its institutions. *New Directions for Evaluation*, 2009(123), 67-85. doi:10.1002/ev.306
- Stopher, P., & Bullock, P. (2003). Travel behaviour modification: a critical appraisal. Presented at the 26th Australasian Transportation Research Forum, Wellington.