Paradigm shifts in transport: scepticism, denial, frustration and (eventually?) implementation in relation to travel behaviour change

Ian Ker

CATALYST (Consulting in Applied Transport, Access and Land use sYSTems)
92 Vincent Street, Mt Lawley, Western Australia, 6050
catalystian@netscape.net

Abstract:
Achieving sustainable mobility requires change to the reasons for travel and the ways in which we travel: where, when, why and how.

Voluntary Travel Behaviour Change consistently delivers successful outcomes valued by the community well in excess of the costs of implementation. These have been well-documented, so why the requirements for ‘pilot/demonstration projects’ as a pre-requisite for mainstream programs that rarely seem to eventuate?

Why the arguments about details of statistical proof – tests that conventional transport planning paradigms would often fail? Why the published reports and papers that ignore the evidence and in some cases present erroneous data? Why is it so difficult to correct the public record when such errors are published? Why are errors perpetuated even when corrections and rebuttals are published – do we need a new paradigm for publication of professional and technical research?

There appear to be disparate standards of ‘proof’ required for different approaches to sustainable mobility. Some are simply seen to be a ‘good thing’ with little supporting evidence; others are made to jump through complex hoops – and, even when they do so, still may not be accepted.

This paper documents and analyses some issues of travel behaviour change acceptance and non-acceptance, using Individualised Marketing as a case study, with reference to both behavioural theory and practical experience.

This phenomenon is not unique to travel behaviour change, but is so pervasive in this field that it is important to understand why it happens and, to the extent possible, correct the public record.
1. Introduction

Marketers and political propagandists know that repetition creates familiarity and a way into long-term memory that is likely to establish an implicit acceptance (http://changingminds.org/principles/repetition.htm). So if we see or hear the same message multiple times, we have a natural tendency to accept it, which is most unfortunate if the message is erroneous.

This could be a description of experience with so-called ‘critical analysis’ of voluntary travel behaviour change (VTBC), which is ironic given that VTBC (or TravelSmart, as it was originally conceived and developed in Perth, Western Australia) has provided more substantial documentation of methods and outcomes (see http://www.transport.wa.gov.au/14974.asp#household for Perth, Western Australia, alone) than almost any other transport initiative known to the present author.

This paper documents some experience with review and interpretation of voluntary travel behaviour change, using Individualised Marketing (the pioneer technique first applied in Perth, Western Australia, in 1997 as a case study, discusses the nature of the research review and publication failures they illustrate and suggests some ways of overcoming such failures.

2. A little history

In 2003, a number of statements to the effect that VTBC initiatives (TravelSmart/IndiMark™, in particular) had not been proven to be effective were published (Stopher and Bullock, 2003). Despite substantial rebuttal (Roth et al, 2003; Ker et al, 2003) in the same venue (26th Australasian Transport Research Forum), these erroneous claims have developed a life of their own, independent of the rebuttal. Subsequent, more positive, papers on travel behaviour change, albeit referring to different intervention techniques (eg Stopher et al, 2009), have not corrected the 2003 criticisms.

Two years later, Paul Mees presented a paper (Morton and Mees, 2005) to the 28th ATRF in Sydney, New South Wales. This paper was criticised from the floor of the conference by the present author on the grounds that it was selective in its use of published evidence, was highly simplistic in its analysis and reflected only one side of the debate that had been held two years before at the 26th ATRF.

In 2010, the identical paper was published in World Transport Policy and Practice (Morton and Mees, 20101), without any attempt to address previous criticisms or reflect published developments and experience since the 2005 paper was written (eg, AGO, 2005; Brög and Ker, 2009).

In 2008, a report on approaches to congestion management (BITRE, 2008) dismissed voluntary travel behaviour change as ‘unproven in their efficacy and cost effectiveness’ (p26) in less than one page – on the basis of only three references, including Morton and Mees (2005), and without reference to a comprehensive review undertaken by another Australian Government agency (AGO, 2005).

Also in 2008, a so-called ‘critical assessment and meta-analysis of empirical evidence’ on voluntary travel behaviour change (Moser and Bamberg, 2008) was published with so many data errors that its conclusions cannot be taken seriously.

These episodes have in common a clear failure of peer review and that they have developed a life of their own with no linkage to rebuttals.

---

1 Since the 2005 and 2010 papers are identical, reference to Morton and Mees (2010) should be taken to include the 2005 paper. Page references are to the 2010 paper.
3. That is not logical

There is an emerging literature on how and why apparently-logical people make illogical decisions (Brafman and Brafman, 2008) or why outcomes may not accord with what conventional wisdom says will happen (Levitt and Dubner, 2010). These are likely to be more immediately important than the arguments about complex systems, indeterminate outcomes and tipping points discussed by Chambers and Ker (1997).

The reasons for apparent illogical behaviour include:

- **Diagnosis bias** based on labeling or stereotyping of the source of information and an inability to re-evaluate an initial diagnosis of a person or situation;
- **Commitment** (individually or collectively as a profession) to a course of action, including previous statements or actions;
- **Loss aversion** – going to great lengths to avoid perceived personal losses rather than maximising personal or community gains; and
- **The chameleon effect** – taking on characteristics that have been, often arbitrarily, assigned to us (Brafman and Brafman, 2008).

All of these result in a common syndrome of ‘dismissing objective data when the information doesn’t fit what they want to see’ (Brafman and Brafman, 2008, p77).

There is also the ‘experts say it is so’ syndrome, highlighted by experience with testing seat belts for children against child safety seats, including the near-impossibility of finding a crash-test laboratory that would test the hypothesis that child safety seats offered no safety advantage over conventional seat belts for children over 2 years old. This was despite fatality data that strongly indicated that there was no significant difference in the fatality rates between the two (Levitt and Dubner, 2010, pp150-158).

The German philosopher, Arthur Schopenhauer (1788-1860), described the experience that new ideas historically have had to endure in the following way: “All truth passes through three stages. First, it is ridiculed. Second, it is violently opposed. Third, it is accepted as being self-evident.”

The factors outlined above go a long way towards explaining why so many commentators on travel behaviour change seem to be stuck at the second stage.

4. Standards of proof

Like Stopher and Bullock (2003), Morton and Mees (2010) emphasise the requirements for statistical validity, ignoring:

- The cumulative sample represented by multiple interventions in Perth, Western Australia – reported results are now for 143,000 persons across 8 projects;
- Repetition and consistency across multiple projects; and
- Demonstrated 99% statistical confidence of behaviour change for the large-scale Brisbane North project (Socialdata, 2007).

Critics focus on the probability of the true result being lower than the measured result as a result of sampling errors. In fact, there is an equal probability that the true result is greater than the measured result. Brög and Ker (2009, pp105-107) show that, for the 13.4% measured reduction in car driver trips (with 99% probability of being greater than zero) in Brisbane North, there is:

- 90% probability that the true result is greater than 6.4%; and
- 80% probability that the true result is between 6.4% and 20.4%.
Put another way, there is only a 10% probability that the true result was less than 6.4% reduction in car driver trips.

This highlights a significant issue for public policy research and its influence. In complex systems it is very difficult to isolate cause-effect relationships, even using control groups, to the same extent that can be achieved in, say, medical research. We should question whether the statistical ‘requirements’ (principally 95% confidence) are appropriate given that an intervention is in areas of public policy where failure does not result in unacceptable consequences. In simple terms, should we deny ourselves a 90% possibility of achievement because of a 10% possibility that we will not achieve what we set out to?

This problem of attribution is not unique. It is a common issue in the evaluation of social policy, especially in an environment of multiple interventions, where, at best, the evaluator has “a variety of quasi-experimental or comparative studies in which…imperfections are acknowledged and in effect adhere to the ‘indicative rather than definitive’ principle…” (Saunders, 2011, p99). In these circumstances, unless we are to throw out the possibility of meaningful evaluation, we need to work with new metaphors:

• Courtrooms not laboratories, in which available knowledge (acknowledging imperfections) provides the basis for establishing what it is reasonable to conclude;
• Indicative and evocative rather than definitive with one-dimensional causality, using the best possible research designs to establish what might be called ‘provisional stability’;
• Alignment rather than attribution, reducing the 'need' for attributional certainty by identifying patterns of evidence that indicate the contribution made by an intervention strategy (Saunders, 2011).

5. Disparate standards for analysis and proof

It is worth noting an apparent double-standard in the quality of research and analysis regarded as acceptable by those who question the statistical validity of the measurement of voluntary travel behaviour change. TravelSmart is consistently ‘required’, by critics, to demonstrate that its measurements and the statistical interpretation of its measured achievements are valid and meet theoretical criteria, but critics do not always apply the same standards to their own data and analysis. For example:

• Stopher and Bullock (2003) misinterpret reported TravelSmart achievements because of a misunderstanding of how results were measured (Roth, et al, 2003), an error that could have been avoided by talking with TravelSmart practitioners. The authors have neither retracted nor corrected their assertions and the errors continue to be disseminated (Morton and Mees, 2010). Likewise, Morton and Mees did not seek information or clarification from those involved with the principal work that they criticise.
• Möser and Bamberg (2008), despite making comments about “weak quasi-experimental evidence” (p19) report results of statistical tests on that same evidence and, even more disturbing, use data (their Table A1) that contain so many errors that the specific conclusions based on theoretical statistical tests are meaningless (Wall, et al, 2011).

Furthermore, the ‘analysis’ of both Stopher and Bullock (2003) and Morton and Mees (2010) is simplistic and superficial, using only selective reported data and, implicitly, starting from the premise that the TravelSmart literature must be in error. These matters have been addressed in a response to Morton and Mees (Ker, 2011) and only some key points are outlined here.

- The single focus on public transport ignores 80% of the car driver trip reduction, where trips were converted to walking, cycling and car passenger.
- Even within the public transport focus, they failed to allow for substantial variations between years (especially the number of school and work days) that render monthly comparisons of patronage unreliable. Conversion to a daily basis (Figure 1) invalidates their (and Stopher and Bullock’s) argument that the increase in public transport solely pre-dated the TravelSmart application in South Perth.

**Figure 1** Change in daily (work/school day) public transport patronage from same month previous year

- Their focus on work trips ignores over 90% of the car driver trip reduction, which was largest for leisure, escort (serve-passenger) and shopping trips.

**Figure 2** Car driver trips before and after IndiMark: South Perth (rounded to nearest percentile of ‘before’ car trips)
• Morton and Mees use Census journey-to-work data to cast doubt on reported results from the South Perth IndiMark/TravelSmart, but ignore extensive change of work trips from car driver to car passenger that was ‘lost’ in their use of a single ‘car’ category.
• Morton and Mees (2010, pp8/9) appear to have a problem with the morality of changing people’s perceptions of reality. Normally, not being a fan of social engineering, I would have some sympathy with this point of view, but in the case of transport we know that people typically:
  o Under-estimate the time and cost of travel by car, and
  o Over-estimate the time and cost of travel by public transport, often through a lack of familiarity with routes, schedules and fare structures (Figure 3).

Figure 3  User perceptions of travel time and cost by car and public transport

This is precisely where one of the sources quoted by Morton and Mees (2010) states that voluntary travel behaviour change can work really well:

“…for public transport, where services and travel quality [are] much higher than perceived, personalised approaches can have very large effects, but where such a gap does not exist the travel behaviour effects could be negligible…” (DfT, 2002, 8.3).

• When they do find evidence of travel behaviour change, it is a reduction in public transport trips in a control group and, without further ado, they pronounce that: “the influence of trends or seasonal influences on transport mode at the population level over a period as short as six months can be discounted” (p16). In fact, this was exactly what the reduction was due to, as investigation by the Victorian Department of Infrastructure (the client for the Alemein project) showed that it was related to the timing of school holidays and, hence, would have had similar effects in both the control group and the project itself. It is precisely the function of control groups to indentify and estimate the effects of such factors.

• Morton and Mees seek to discredit results by saying that survey respondents might consciously or sub-consciously seek to comply with implicit expectations of the intervention (p12), including the ‘good subject’ effect (p17) with falsified responses or non-responses (the latter, presumably, in the case of those who had ‘failed’ to change behaviour but invoked the apocryphal George Washington response of being ‘unable to tell a lie’).
Paradigm Shifts in Transport: Scepticism, Denial, Frustration and (Eventually?) Implementation in Relation to Travel Behaviour Change

There is no evidence of this. Indeed, it could only be even potentially significant in the case of a panel survey or where the surveys were not differentiated from the intervention. Neither of these applies to TravelSmart/IndiMark.

Morton and Mees also invoke a strange ‘doppelgänger’ of the expectancy effect, this time in the control group (p17), when they suggest that some in the control group, knowing that TravelSmart is under way in the area (itself not quite true, as control groups are external to the area in which TravelSmart is undertaken), might “under-report their use of public transport, walking and cycling and/or over-report their car use to ensure that the researchers got the comparative result they wanted” [emphasis added by current author].

6. Reporting research: responses, rebuttal and retraction

At the 26th ATRF, Stopher and Bullock (2003) was subject to a written rebuttal and presentation (Roth et al, 2003; Ker et al, 2003). Whilst the rebuttal was published electronically with the ATRF papers (http://www.patrec.org/atrf.aspx), there is no linkage between the original paper and the rebuttal. Those who find reference in other places to the Stopher and Bullock paper and then access it on the web are given no indication that there was dissent from its content or conclusions.

The original paper has taken on a life of its own, without qualification or contest.

At the 28th ATRF, Morton and Mees (2005) was subject to dissent and rebuttal from the floor by the present author, but there was no mechanism for this to be published in the proceedings. Even substantial rebuttal from the floor was only possible because the present author had been alerted to the paper and had been able to get an advance copy in order to prepare an informed response. Unless conference attendees are able to get papers in advance (as was the case two decades ago), it is very difficult for them to make informed and useful comment – in effect, professional debate is stifled by the common practice of presenting a CD of papers to delegates when they arrive at the conference rather than distributing the papers in advance.

In 2008, Möser and Bamberg (2008) published what purported to be a ‘critical assessment and meta-analysis of empirical evidence’ on soft policy measures for transport. The data they used for household-based travel behaviour change contained so many errors, some of which were so obvious as to be clear simply on the basis of professional knowledge without referring to the source documents, that their conclusions cannot be taken seriously (Wall et al, 2011). It took repeated attempts over three years by the authors of the response for the Journal of Environmental Psychology to agree to publish this response.

Again, the original article has taken on a life of its own, without qualification or contest. In a recent project on active transport, the present author had to intervene to correct the use of Möser and Bamberg’s specific conclusions in the following terms:

“Möser and Bamberg’s data on community interventions include a large number of errors that make the specific conclusions invalid despite the stated statistical test (Wall, et al, 2011). The results from one specific intervention methodology (Individualised Marketing) have been shown to be consistently between 4% and 14% reduction in car driver trips (Ker, 2008).

7. Research on congestion

Closer to home, a report on approaches to congestion management (BITRE, 2008) dismissed voluntary travel behaviour change as ‘unproven in their efficacy and cost effectiveness’ (p26) in less than one page – on the basis of only three references, including Morton and Mees (2005), and without reference to a comprehensive review
undertaken by another Australian Government agency, and available at the time, which concluded:

“For households, predicting the effectiveness of a Travelsmart project is now essentially a solved problem...While individual outcomes vary with geographic location, what can be said broadly is that community-based household projects will achieve a reduction in car travel of 5–15%, and this change appears to be sustained for several years without further intervention.” (AGO, 2005, p53)

It took over a year for BITRE to add, under pressure, a qualification to this report – but only on its website – drawing attention “firstly, to some larger-scale current initiatives of Australian jurisdictions and, secondly, to more recent literature on the effectiveness of the travel behaviour change approach in relation to its various objectives, particularly reducing motorised trips and reducing greenhouse gas emissions” and providing a large number of links to other relevant documentation (http://www.btre.gov.au/info.aspx?ResourceId=680&NodeId=59).

But the report itself remains dismissively the same, despite much of the ‘more recent literature’ referred to on the website actually having been available at the time the BITRE report’s being written. Once downloaded, the report still makes the erroneous assertions about travel behaviour change being ‘not proven’, without qualification.

Moreover, the report states, on the basis of no declared evidence, that: “Irrespective of the level of effectiveness of such approaches in altering travel patterns, it is clear that it is a costly strategy” (BITRE, 2007, p26). In fact, evaluations of household TravelSmart have consistently shown benefit-cost outcomes that could be reduced by a factor of ten and still compare favourably with conventional urban transport projects (see, eg, Ker, 2004).

The BITRE effectively states that more research is needed. This is in direct conflict with the statement by the Australian Greenhouse Office:

“We advise the State TDM Managers and senior decision-makers that further large-scale evaluation of household projects is not a good return on the very large investment involved. The one exception would be long-term tracking—such as the WA Department of Planning and Infrastructure has been conducting in South Perth.” (AGO, 2005, p53).

Not only does a requirement for project by project evaluation add to the cost of TravelSmart and, hence, reduce the reach of any level of funding, it also tends to push decision-makers towards ‘pilot projects’ that are generally small-scale. It is, consequently, difficult to meet the requirements for statistical validity – a vicious circle, made worse the fact that the evidence shows that the scale of household-based initiatives has a clear (and positive) effect on the outcome in terms of car trip reductions (Figure 4).

Figure 4  Travel behaviour change outcomes and scale of intervention

Source: Brög and Ker (2009, p86)
8. Being overtaken by events

This work itself (BITRE, 2008) is a response to the escalating cost of congestion identified in earlier BITRE work (BTRE, 2007). This report provides comprehensive and robust estimates of avoidable congestion costs in Australian capital cities. The estimates are provided relative to traffic volume and by year to 2020, which allows incremental costs to be derived on either basis. Incremental congestion costs are, almost by definition, higher than average costs and increase more rapidly (see Figure 5).

Figure 5  Average and incremental congestion costs: Sydney, Brisbane and Perth

These values are particularly useful when evaluating non-road initiatives, for which access is not always available to traffic models that produce estimates of (a) total vehicle travel and (b) travel times and vehicle operating costs, which are the principle components of congestion cost. In the absence of estimates of overall traffic volumes, there is little option but to work on the year-by-year values.

There is however, a flaw – and it is buried literally at the end of the report. The traffic volume estimates for future years are largely based on a US oil price of less than US$55 a barrel (Figure 6).

Figure 6  BTRE oil price assumptions for traffic forecasting

In February 2011, oil is US$85/barrel, but the pump price of petrol has been moderated by the increasing value of the Australian dollar versus the US dollar and the depressing effect of the ‘global financial crisis’ on the oil price itself.
In 2010, the average price of oil was US$71.20/barrel (West Texas Intermediate - WTI) but inflation and the appreciating Australian dollar reduced the real (Australian) price of oil by 7% and the price of petrol at the pump by 9%. At the beginning of May 2011, the price of oil was US$112.80 (WTI) and the Australian dollar was worth US$1.09 – a real Australian oil price that is 20% higher than in 2007, despite a further 10% appreciation of the Australian dollar since December 2010.

The generally-conservative International Energy Agency is forecasting that oil prices will exceed $100 a barrel in the next five years and by 2035 will exceed $200 a barrel (IEA, 2010). It is also stating that oil prices as high as $200 per barrel will be needed to make non-conventional sources of oil (such as oil shale) economically viable (http://www.liveoilprices.co.uk/oil/peak_oil/11/2010/iea-forecasts-that-peak-oil-production-started-in-2006.html).

Even with the protection of the appreciating Australian dollar, the 2011 traffic and congestion estimates will be high because the real price of petrol in April 2011 was already 2% higher than the average for 2007. If road traffic growth to 2020 is, say, half that estimated by BTRE, the incremental cost of congestion will be 15% lower in 2020 than the BTRE estimates. Using rule-of-thumb price elasticity of demand estimates of -0.2 to -0.3, this requires only 30% to 50% increase in the effective real price of petrol (ie after allowing for fuel efficiency improvements). A doubling of the effective real price of petrol by 2020 would keep incremental congestion costs at 2011 levels (30-35% lower than BTRE estimates for 2020).

Whether or not the BTRE assumption about future oil price was reasonable at the time, circumstances have clearly changed. Indeed, it would be surprising if they had not, given the demonstrated volatility of oil prices at that time. It would have been prudent to undertake sensitivity analysis for traffic forecasts and congestion costs for higher oil prices.

More important, now that oil prices are twice those assumed in the BTRE’s congestion cost analysis, it would be highly beneficial to revise the congestion cost values and publish an addendum to the 2008 report.

9. Peer review

All of the examples above were subject to some form of review before they were published. Despite this, they emerged complete with errors of assumption, fact or analysis that a competent and knowledgeable reviewer should have identified.

• Did the reviewers not comment on those errors?
• Did the authors refuse to amend papers in the light of comments made by reviewers?
• Was publication authorised despite errors that were pointed out but not responded to?²

In one case, Morton and Mees (2005), one referee review strongly recommended that the authors be required to check many of their statements and conclusions with those familiar with IndiMark™ as the reviewer did not have sufficient direct knowledge to be able to comment on their validity. However, this was not done, despite neither Morton and Mees (in 2005 or between 2005 and 2010) nor Stopher and Bulloch (in 2003) having contacted any of the people who had actually been involved in the development, delivery or evaluation of the initiatives of which they were so critical.

² Crude oil prices were as low as $16.80/barrel at the end of 2001, exceeded US$126/barrel in June 2008 and went below US$40/barrel just 7 months later, in January 2009.

³ It is pleasing to note that the organiser of the 2011 ATRF required authors to document their responses to reviewers’ comments.
Whatever the reason, it seems that the peer review process has been ineffectual in the cases discussed in this paper. This is not an observation confined to transport research; problems with review and publication of research (both peer review and the quality of the research itself) are so widespread that the UK Research Integrity Office has found it necessary to publish guidance on peer reviewing in its Code of Practice for Research (UKRIO, 2009) and an Information Note on retractions in academic journals (UKRIO and COPE, 2010).

An assessment of peer review in the field of medical research (Smith, 2006) concluded:

...peer review is a flawed process, full of easily identified defects with little evidence that it works. Nevertheless, it is likely to remain central to science and journals because there is no obvious alternative, and scientists and editors have a continuing belief in peer review. How odd that science should be rooted in belief. (Smith, 2006, p182).

Smith’s acceptance that there is no obvious alternative might have been understandable in 2006 (or, more likely, 2005, since he observes that: many journals, even in the age of the internet, take more than a year to review and publish a paper). In the last five years, however, there has been an explosion of interactive capabilities on the internet that could provide the basis for a re-engineering of review, publication and criticism processes.

10. A 21st century paradigm for publishing research

It is ironic that in the age of ‘instant communication’ it should be so difficult to review research publication effectively or to correct erroneous information or update analysis and conclusions when key assumptions of or inputs to published research are found to be incorrect.

In these days of electronic publishing of reports and journals, it should be a simple matter to update documents to remove errors and omissions and even generate discussion about the subject of the paper. The British Medical Journal, as well as publishing selected responses to articles in conventional form, provides a ‘rapid-response’ facility on the journal’s website, so that readers can immediately see differences of opinion or countervailing arguments. These rapid-responses are clearly visible at the time of reading the article on-screen.

Moreover, where there is a known and documented difference of view, the two are also clearly-linked in a very visible way (see, eg, http://www.bmj.com/content/342/bmj.d2168.full and http://www.bmj.com/content/342/bmj.d2177.full).

Some other medical journals provide citation lists for articles (eg the Journal of the American Medical Association – see, eg, http://jama.ama-assn.org/content/300/22/2631.full), which would at least provide the opportunity for rebuttals to be listed integrally with the original article.

However, transport journals (and conference proceedings) generally simply present hard-copy in electronic form. Organisations publishing research reports have essentially done the same.

A 21st century paradigm for reporting research would not treat published work as an end in itself but as a generator of debate, discussion and further research, including application of the research. Electronic documentation can be readily updated to reflect criticism and commentary, just as computer software is updated to respond to problems, so that future readers are able to benefit from the debate rather than being left to make the same mistakes over again.
11. Conclusion

It is important to have informed debate and criticism of published research, but experience with voluntary travel behaviour change suggests that much criticism is poorly-informed and often does not meet the standards demanded of the research by the critics themselves.

As transport research and policy become more complex and moves into areas where traditional methods may be inadequate, it is essential to move publication itself into faster and more inter-active modes. Modern information and communications technologies provide a means by which research and criticism can be both faster and better informed.

More fundamentally, electronic documentation can be readily updated to reflect such criticism and commentary so that future readers are able to benefit from the debate.

References


Ker I (2004). TravelSmart Individualised Marketing: Evaluation. ARRB Transport Research FOR Deparment for Planning and Infrastructure: Perth, WA.


