Causaal onderzoek op de relatie
tussen mobiliteit en ruimte

Paul van de Coevering
NHTV Breda
Coevering.p@nhtv.nl/

Kees Maat
TU Delft
c.maat@tudelft.nl

Bert van Wee
TU Delft
g.p.vanwee@tudelft.nl

Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk
20 en 21 november 2014, Eindhoven

Samenvatting

Tot op heden is een groot deel van de kennis op de relatie tussen mobiliteit en ruimte gebaseerd op cross-sectie onderzoek. Hierbij wordt op één moment in de tijd data verzameld. Voordelen zijn onder meer de relatief eenvoudige toepassing en de snelle beschikbaarheid van de resultaten. Met deze onderzoeken zijn statistische verbanden aangetoond tussen mobiliteit en ruimte (de dichtheid van woonwijken hangt bijvoorbeeld samen met het autogebruik). Dit levert echter nog geen afdoende bewijs voor het bestaan van een causale relatie. Voor het aantonen van causaliteit moet onder meer een oorzaak gevolg relatie worden vastgesteld. De vraag is of een verandering van de ruimte ook daadwerkelijk leidt tot een verandering van de mobiliteit. Om dit te onderzoeken zijn onderzoeksdesigns nodig met meerdere meetmomenten in de tijd. Deze paper geeft een overzicht van deze designs en beschrijft aan de hand van empirische studies hoe ze kunnen worden toegepast bij onderzoek op de relatie tussen mobiliteit en ruimte.

De keuze van het onderzoeksdesign zal in de praktijk afhangen van de onderzoeksvraag (causaal of associatief) en de praktische mogelijkheden en beperkingen (geld, tijd etc.). Natuurlijke experimenten – met een meting voor en na een verandering in de context - vormen de meest robuuste onderbouwing voor causaliteit. Dit design wordt aanbevolen om de effecten van veranderingen in de fysieke omgeving zoals de aanleg van nieuwe fietspaden of de komst van een nieuw winkelcentrum te bepalen. Prospectieve longitudinale designs worden aanbevolen wanneer ruimtelijke veranderingen worden onderzocht die het gevolg zijn van een bewust keuzeproces van mensen zoals een verhuizing of een baanverandering. Hiernaast kunnen interactie-effecten tussen de ruimtelijke context en veranderingen in de huishoudensomstandigheden worden onderzocht. Zijn mensen bijvoorbeeld vatbaarder voor (mobiliteits)gedragsexperimenten na een verandering van de ruimtelijke context?

Afhankelijk van de beschikbare middelen zullen onderzoekers in de praktijk soms moeten terugvallen op retrospectieve designs of herhaalde cross sectie metingen. De voor- en nadelen van deze designs op het gebied van de causaliteit en de praktische toepasbaarheid komen uitgebreid aan bod.

Trefwoorden: onderzoeksdesign, longitudinaal, causaliteit, attitudes, mobiliteit, ruimte

1 Introduction

For decades, researchers have recognised the impact of the built environment (BE) on travel behaviour (TB) (Maat & Timmermans, 2009; Van Wee, 2011). Diverse policy measures, such as densification, mixed use development and the improvement of cycle routes, have been undertaken to provide viable alternatives to car use and to alleviate problems such as air pollution and road congestion. Often, these measures involve high costs and/or have long-lasting implications. Reliable evidence on the causal effects of these BE changes on TB is therefore important. To date, the vast majority of empirical studies on the BE-TB link applied cross-sectional designs that provide a snapshot of the variables at a single moment in time. Although these studies provide valuable insights, they cannot infer causal relationships because they (Kitamura, 1990; Mokhtarian & Cao, 2008):

- do not allow the assessment of the impact of BE changes on TB over time;
- are vulnerable to third (confounding) variable influences;
- neglect the dynamics involving behavioural change which include time lapse and (temporal) mismatches between circumstances and behaviour.

To overcome the limitations of cross-sectional studies, the use of research designs that provide data for two or more moments in time (hereafter referred to as "multi-period designs") is increasingly being advocated in literature (see Mokhtarian & Cao, 2008). As yet, few studies have been based on these designs, partly because the practical application of these multi-period designs (e.g. repeated cross-sections, longitudinal and experimental designs) is generally considered more complex, expensive and time consuming (Transportation Research Board, 2009). However, there is no comprehensive overview that describes how and to what extent multi-period designs can be applied to the BE-TB link. Therefore, the extent and nature of the advantages and disadvantages that arise from applying the variety of designs remains unclear.

This paper builds on the previous work and aims to give an overview of how multi-period designs can be applied in research on the BE-TB link. It therefore provides a classification of research designs. Methodological advantages and disadvantages with regard to causality and data collection are discussed. Empirical studies will be used to illustrate how these designs can be applied in practice in studying the BE-TB link. This will aid researchers to select an appropriate design and deal with common issues in future work in this field.

Section two recaps the causality debate that surrounds the BE-TB link and describes a conceptual model. Section three describes the methodological advantages and disadvantages of a classification of multi-period research designs and describes how these designs can be applied to the BE-TB link. The last section summarises the findings and defines implications for future research on the BE-TB link.

2 Conceptual framework and limitations of cross-sectional designs

Figure 1 conceptualises the relationships between BE-TB and major third variables that are often included in studies in this field. Earlier studies have suggested a direct relationship between the BE characteristics of residential locations and TB; for example, living near a railway station would encourage travelling by train [link 8]. Most of these studies accounted for the influence of objective socio-economic and demographic variables on TB [link 3]. Lately, subjective variables such as attitudes to TB have entered the equation [link 7]. Attitudes are considered to be the product of individual and household characteristics, including the phase in the life cycle and lifestyle preferences [link 1].

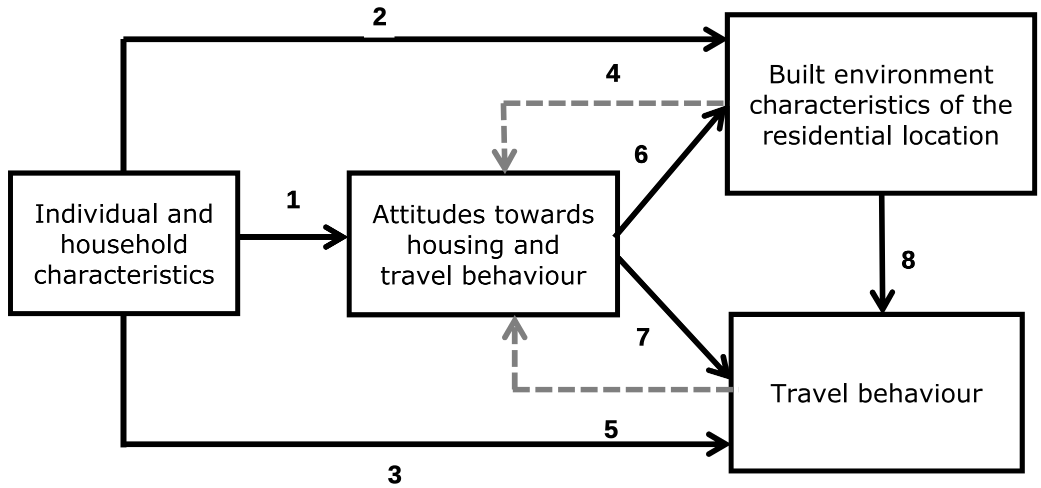

FIGURE 1: Conceptual model of relationships between BE-TB and intervening variables. Adjusted from Bohte (2010)

The question here is whether the BE influences TB directly or whether this effect occurs through residential choice, ceteris paribus. The residential self-selection (RSS) hypothesis suggests that household location choices are based on travel-related attitudes and socio-economic and demographic characteristics [links 2, 6] (Cao *et al*, 2009). For instance, someone who prefers to walk may settle in a neighbourhood that is conducive to walking. In this case, it is not the BE that causes someone to walk, but rather it is his or her pre-existing positive attitude towards walking that causes that person to choose a walkable BE, which makes more walking possible. However, academic literature also suggests two-way causation between attitudes, residential location choice and TB [links 4 and 5]. People may alter their attitudes to make them 'fit in' with their current residential and transportation choices. through the process of "cognitive dissonance" (Festinger, 1957). Furthermore, attitudes may change over time due to changes in the BE or changes in the household composition (Chatman, 2009). For example, after living in a residential location close to a railway station for some time, individuals may change their attitudes and start to appreciate public transport (Bohte, 2010).

To date, most studies have found significant associations, ceteris paribus: residents in neighbourhoods with higher densities, mixed use developments, short distances to destinations and good facilities for public transport, cycling and pedestrians, tend to drive less and make more use of alternative modes of transportation (Ewing & Cervero, 2010). However, the majority of studies found a weakening of the effect of the BE on TB after controlling for RSS (Transportation Research Board, 2009; Cao et al, 2009). Furthermore, most previous studies have only considered the influence of attitudes on behaviour and have not taken into account the possibility that people may adjust their attitudes to their current BE (Bohte, 2010). Hence the direction and nature of causality on the links between BE-TB and attitudinal characteristics remain uncertain. Are positive attitudes towards public transport in a BE with a high level of public transport provision the result of RSS? Or are these positive attitudes the result of the BE that is conducive to using public transport?

To disentangle the sequence of events and provide evidence for causality, four conditions should be met (Shadish, 2002; Singleton & Straits, 2005):

1. association (significant statistical relationship between two variables);
2. time precedence (the cause precedes the effect);

3. non-spuriousness (the relationship between two variables cannot be attributed to another variables);
4. plausibility (there should be a logical causal mechanism for the cause and effect relationship).

Although cross-sectional studies provide ample evidence for the associations, they fail to meet the other three conditions for causal inference (Mokhtarian & Cao, 2008). In general, multi-period designs provide more opportunities for establishing a causal link. However, there is a wide diversity in types of multi-period designs which come with various advantages and disadvantages regarding causality and data collection. The basic differences between these designs and opportunities to apply them in studies on the BE-TB link are discussed in the next section.

3 Applying multi-period designs on the BE-TB link

Table 1 classifies the main research designs according to their ability to support the criteria for causal inference. The classification will be used as a framework for the discussion on applying these designs. For each design we will discuss its strengths and its potential limitations regarding causal inference, opportunities for implementation, and issues that commonly arise during implementation. Examples of studies into the BE-TB link and of studies on adjacent research fields of physical activity, environmental psychology and transportation are included to illustrate how research designs can be applied.

3.1 Experimental Designs

The randomised experiment (the classic before-after random-assignment control group design) is often considered the "gold standard" for inferring causal effects. Researchers exert full administrative control over the interventions and randomly allocate participants to experimental and control groups. Consequently, unobserved variables that may influence the outcome of the study will be evenly distributed between the groups. This minimises the threat of third variable bias (Shadish *et al*., 2002).The effects of the intervention can be assessed by testing the differences in the outcome variables between the research groups before and after the intervention. Additionally multivariate analyses can be applied to examine the causally mediating mechanisms of these effects. *Quasi-experimental designs* share many characteristics and are often used in situations where full randomisation is not possible, for instance due to the context of the study and/or ethical or political considerations. Consequently, the intervention is exogenously assigned or randomly assigned at a higher level of grouping (e.g. to schools, departments, neighbourhoods etc.) (Remler & Van Ryzin, 2010).

Applying an experimental design directly to the BE-TB link would mean that researchers would (randomly) assign residents to different residential areas, or subject them (randomly) to changes in the BE. Aside from the financial costs, it would of course be impractical and unethical to move households, just for the sake of a study (Mokhtarian & Cao, 2008). Nonetheless, experimental designs can provide valuable information about causality when characteristics of the BE are incorporated in the evaluation of travel behaviour change experiment such as awareness and promotional campaigns. How large of a role does the BE play in the success of these experiments? Riley et al. (2012) for instance analysed how neighbourhood walkability influenced the effect of a randomised intervention aimed at increasing moderate and vigorous physical activity. They compared the results of the intervention in neighbourhoods with high and low levels of walkability, but found no significant interaction effects between walkability and the intervention. Furthermore, changes in the BE (due to a move or spatial or infrastructural measures) are alleged to provide a unique context for these behavioural experiments as they motivate people to reconsider their TB and make them more receptive to the experiment (Bamberg, 2006; Graham-Rowe, 2011). Bamberg (2006) used a residential move to Stuttgart as the context for an experiment and allocated movers to an experimental group – that received an intervention in the form of a free public transport

TABLE 1 Classification of Study Designs

			Causality				Data collection and analysis					
	Number of waves	Design	Association	Time precedence	Causal mechanism	Third variable	Required Sample size	Initial non-response	Practical issues	Drop-out	Decline of accuracy	Periodic effects
Experimental design	≥ 2	Randomised experiment	++	++	++	++	++	-	--	Yes	Yes	Yes
	≥ 2	Quasi experiment	++	++	++	+	+	-	-	Yes	Yes	Yes
Observational design s	≥ 2	Natural experiment	++	++	++	+	+	-	-	Yes	Yes	Yes
	≥ 2	Prospective longitudinal studies	++	++	++	0	+	-	0	Yes	Yes	Yes
	1	Retrospective (quasi-) longitudinal studies	++	0	0	0	0	0	+	No	No	Yes
Independent cross-sectional samples	≥ 2	Repeated cross-sections / pseudo panel	++	+	-	-	--	++	+	No	No	Yes
	1	Single cross-section	++	--	--	--	--	++	++	No	No	No

Notes:

- The sign (++, 0 , --) indicates if the design scores relatively positive, negative, or neutral on the criterion. Yes/No means that this is or is not an issue for the research design concerned.
- More detailed classifications in research designs are possible (e.g. based on the number of measurements or the presence of comparison groups) but this is not the aim of this section.

After: Remler & Van Ryzin (2010), Shadish (2002); Kitamura (1990) and Ortúzar (2011).

ticket and specially tailored information – and a control group that received no intervention. He concluded that only in the experimental group the move had actually provoked e a change in TB, indicating that the intervention was necessary as “a last push”. Ideally, these experimental studies should include an additional comparison group of people who are not affected by the BE change. This would have provided more detailed insight into the role of the BE changes and the interaction with the experiment.

For (quasi-)experimental designs, the advantages for causal inference come at the cost of a lower practical applicability. Ideally, two or more follow-up measurements will be taken to determine the short- and longer-term effects as the effects of the interventions may increase, diminish or remain stable over time (Graham-Rowe *et al*., 2011; Bradley, 1997). It goes without saying that designs that require two or more measurements and a controlled assignment of an intervention take more time and effort and are more costly than a single cross-section or retrospective design. (Ortúzar, 2011). Furthermore, designs that require two or more measurements place a higher burden on the respondents. Therefore, initial non-response is generally larger and non-response increases between the research waves as respondents drop out– because they move, die or are simply not willing to participate anymore (Behrens and Del Mistro, 2010; Kitamura, 1990; Shadish, et al., 2002). Also, some caution is required regarding the analysis of the effects as not all effects are attributable to the experiment itself. Changes in reported behaviour over time may also occur due to declines in reporting accuracy and periodic effects. These effects should not be confused with intervention effects (Shadish, 2002; Meurs, 1991). Declines in accuracy occur for instance because respondents become more aware of the issues related to the study. These ‘test effects’ can be minimised by preventing people from associating the evaluation methodology (e.g. questionnaires) with the experiment. Bamberg (2006) for example presented his study as a general project to analyse daily mobility patterns and the promotional materials were sent to the participants by another company that did not make any reference to the research project. On the other hand, respondents may also get bored with answering the same questions repeatedly which may result in an increase of reporting errors (Meurs, 1991). Periodic effects are related to changes in the context such as changes in weather conditions, the economic situation etc. (Shadish, 2002).

3.2 Observational Designs

If the assignment of an intervention is beyond the control of the researchers, they can adopt observational designs that rely on naturally occurring (not planned by researchers) events.

In *Natural experiments*, the experimental condition happens naturally or at least in a way that is not related (exogenous) to the research results. They can be applied to examine the effects of infrastructure improvements on TB (e.g. Goodman *et al*, 2013), for example. It has been recognised that some developments may partially be the effect, rather than the cause of TB changes (for instance due to community lobbying for new bicycle infrastructure). However, most of them can reasonably be considered as exogenous events (Krizek *et al*., 2009). Ideally they use a before and after design with comparison groups. Because researchers do not have control over the assignment, they are often not considered “real” experiments but rather a sophisticated form of observational design (Shadish, 2002). With regard to causality, natural experiments can be comparable in power to quasi-experiments.

To date, only a handful of studies applied this design to the BE-TB link. An explicit assessment of spatial interventions was not found in current literature. Available studies mainly assessed the effects of changes in the provision of public transportation (e.g. Arentze *et al*., 2001; Chatterjee, 2011; Yanez *et al*., 2010). Chatterjee (2011) assessed changes in TB after the introduction of a new bus rapid transit service using one baseline measurement and three follow-up measurements. Results showed that people gradually adapted to the new services, which resulted in an increase in bus use over time. In recent years, the application of this design has progressed on the link between physical activity and the BE (e.g.; Evenson, 2005; Giles-Corti *et al*., 2008; Ogilvie *et al*. 2010; Goodman *et al*., 2013). Goodman *et al*. (2013) conducted baseline, one-year and two-year surveys to examine the TB effects of a new walking and cycling route and found that the infrastructure catered mainly for existing walkers and cyclists.

An important issue for practical application is that researchers have limited control over the intervention and contextual influences. Studies by Yanez (2010) and Evenson *et al*. (2005) were seriously affected by a chaotic and delayed introduction of new transport facilities. To address these issues and maximise opportunities to apply these natural experiments, collaboration between transport and land-use planners and researchers is important (Ogilvie *et al*, 2011; Boarnet, 2011). Furthermore, the selection of proper treatment and comparison groups is far from straightforward: all the aforementioned studies lacked a comparison group; it appears to be difficult to determine the boundaries of the (intervention) areas that are exposed to infrastructure (re)developments; and comparable neighbourhoods and participants are often difficult to find (Ogilvie, 2010; Stopher *et al*., 2009). Suggested alternative approaches to the inclusion of comparison groups are:

- data and information from alternative sources to control for unknown influences; Ben-Elia & Ettema (2011) assessed the impact of rush hour avoidance and used traffic count data from the trajectory being studied to test for disruptions or unexpected changes in the traffic flow;
- a "dose-response" assessment (Ogilvie, 2006); the distance to newly developed infrastructure (or access points like a public transport stop or on-ramp) can be used as a proxy for exposure and incorporated in the effect analysis;
- including multiple baseline and follow-up measurement points to identify a break in TB trends (Merom, 2003; Graham-Rowe *et al*., 2011); however, this increases the risk of drop-out as it requires more commitment from the respondents.

With *prospective longitudinal designs*, researchers look ahead and observe what happens to each of the participants and analyse the effects of these events on the output variable over time. These run for longer periods of time and may have multiple points of measurement. General transportation panels such as the Dutch Mobility Panel (LVO) and the German Mobility Panel (GMP) are typical examples (Ortúzar *et al*., 2011). Other examples include the assessment of the impact of residential moves on TB (Krizek, 2003). One important disadvantage is that the majority of the events are not exogenous because they are based on deliberate decision making and self-selection (Remler & Van Ryzin, 2010). This leaves room for the effects of third variables. However, like the previous designs, prospective studies enable the researchers to focus on changes in the dependent and independent variables between two or more periods in time. This means that unobserved third variables that remain stable over time can be well controlled for, reducing the amount of noise that is often apparent in cross-sectional studies (Handy, 2005). Data collection is a bit easier because researchers do not have to deal with an experimental condition. However, the concerns regarding design with multiple measurements (e.g. drop out) also apply to this design.

There are two approaches to assess the impact of the BE on TB in the occurrence of endogenous (self-selected) changes in household circumstances. The first involves the environmental context explicitly, and assesses the impact of, for instance, residential moves or job location changes. Krizek (2003) examined changes in BE characteristics and household TB between two consecutive years and subdivided a research sample from a general-purpose travel panel survey into an "experimental" group with people who had moved house and a comparison group of people who had not. He found that households changed their TB after relocation. Higher neighbourhood accessibilities in particular reduced the number of miles travelled by car. Meurs and Haaijer (2001) came to comparable results using a two-wave longitudinal dataset based on the Dutch Time Use Study. In recent years, this design has also been applied in research on the link between physical activity and the BE (e.g. Giles-Corti *et al*., 2013).

Under the second approach, the impact of life-course events on people's TB over a longer period of time is assessed. The role of the BE is used as an explanatory factor for the impact of life-course events on TB. Bohnet & Gertz (2010) analysed data from a German multipurpose household panel survey and found that higher density housing increases the probability that residents will postpone the purchase of a car and the acquisition of a driver's license, and later in life it increases the probability that they will sell their car, if they have one. Clarke *et al*. (2009) found that living in more pedestrian-friendly neighbourhoods was associated with a reduced probability of mobility disability in older age.

One of the major concerns of prospective research designs is the endogenous nature of life-course events (including residential relocation). To ensure that the observed relationships on the BE-TB link are not spurious, it is possible to include confounding (third) variables in the analysis. This notwithstanding that control variables are also recommended for more rigorous designs as the influence of third variables cannot be completely eliminated through research design on the BE-TB link. Another issue is drop-out, especially for those people that moved house. Methodologies such as providing incentives and maintaining contact proved effective in reducing these drop-out rates (see Yanez, 2010 and Ortúzar, 2011). Furthermore, these studies often lack a comparison group which reduces their internal validity. Possible alternatives are discussed in the description of the natural experiments. A problem specifically related to the second approach is that major life-course events and subsequent changes in TB do not take place regularly. This means that large time spans are needed to yield significant results, and this can be difficult to achieve with prospective designs (Woldeamanuel *et al*., 2009). Therefore, many studies on life course events are based on retrospective designs.

The *retrospective (quasi-)longitudinal designs* involve one-off surveys of individuals that ask respondents to recall information about events, activities or other phenomena that happened to them in the past (Behrens and Del Mistro, 2010). Research suggests that respondents are able to recall major lifecycle events and report some of their key attributes. However, partial or inaccurate responses are likely to be due to memory errors and other variables, notably attitudes, cannot be recollected reliably retrospect (Verhoeven, 2010). This reduces the ability to meet the time precedence criteria and to disentangle causal mechanisms for these variables. Their ability to infer causality is thus weakened. However, retrospective surveys allow for observations over longer time spans than would be feasible with panel surveys and data can be attained more quickly, whereas panel data is able to cover more and yields more detailed information.

Similar to prospective designs, retrospective designs are applied in two different ways. The first approach assesses the TB effects of specific endogenous (e.g. residential moves) or exogenous (e.g. infrastructural or spatial interventions) changes in the BE. Handy *et al*. (2005) assessed the impact of residential moves retrospectively by classifying respondents into movers – those who moved house within the last year – and non-movers. Current travel-related attitudes were accounted for to reduce the risk of spurious results. They found a significant association between the BE and TB. Cao *et al*. (2007) and Aditjandra *et al*. (2012) used a comparable approach and also found significant associations. However, these studies lacked a control group of non-movers. Boarnet *et al*. (2005) conducted a retrospective assessment of the effects of an exogenous change in the form of improved walking and cycling routes on children's' active travel to school. They asked parents retrospectively about changes in their children's walking and cycling frequency after infrastructure improvements were made. Counter to expectations, children who used these improved routes showed a decrease in walking and cycling. However, this was against a backdrop of an overall decline in walking and cycling rates and the researchers eventually concluded that the children who used the improved facilities were more likely to increase their walking or cycling to school than their counterparts who did not use these facilities.

The second retrospective approach enables researchers to assess the influence of the BE during life course events over a longer period of time. Using this design, Beige & Axhausen (2012) examined the impact of life course events on TB. The retrospective design allowed them to cover a period of 20 years. They used vehicle tool ownership (the ownership of cars and public transport tickets) and general TB indicators like the "most used mode of transport" for commuting trips to assess changes in TB. They concluded that life course events including changes in residential moves and changes in occupation create important opportunities for changes in TB.

A disadvantage of the practical application of this design is that the retrospective questioning of attitudes and specific daily TB is not reliable. Handy *et al*. (2005) therefore used a relatively short timeframe and asked the respondents to indicate changes in their TB retrospectively on a 5-point Likert scale. They only asked about current attitudes and considered these to be stable over the relatively short time period. As the authors acknowledge, this limits the ability to disentangle the causal chain on the BE-TB link. Another issue is that these studies do not always include a comparison group of non-movers or another form of comparison. The aforementioned study of Boarnet *et al*. (2005) illustrates its importance. It is clear that the authors' conclusion – namely that the improvement of facilities has a positive effect on walking and cycling – would have been different if no comparison group had been included in the research.

3.3 Independent Cross-sectional Samples

Independent cross-sectional samples involve taking independent random samples from the same target populations over time. They can be conducted before and after contextual changes take place or at fixed points in time to assess the changes in the average population parameters (trends). Continuous national travel surveys in Germany (KONTIV) and the Netherlands (NTS) are typical examples (Ortúzar *et al*. (2011). A major advantage is that they can be undertaken over longer periods of time because the independent samples are not affected by drop-out. A major disadvantage is that they do not measure the same respondents across time, making it impossible to identify or explain intra-personal change or to resolve issues of causal sequence.
However, repeated cross sections can be used to construct *pseudo-panels* that track cohorts of individuals or households with common characteristics (data of birth, gender, etc.) over time. In other words, in a pseudo-panel, it is not individuals that are tracked over time, but rather groups of individuals. The average group characteristics are treated as individual observations. This design has been frequently used in transport surveys in recent years (Weis & Axhausen, 2009).

Firstly, repeated cross-sections are used in studies to estimate aggregate changes before and after specific BE changes. Lovejoy *et al*. (2013) surveyed two samples of respondents, one the year before and one the year after the opening of a big box retail store. They asked them to recall the characteristics of their most recent shopping trip and their general shopping behaviour in the last year. The study reveals that trips to the new retail store had largely displaced vehicle trips to other more distant retail facilities resulting in an overall reduction in the number of vehicle miles travelled for shopping. Secondly, repeated cross sections from national household transportation surveys or census statistics are often used to detect TB trends in the population. We refer to Ortúzar *et al*. (2011) for an extensive review of continuous transportation surveys. Because cohorts for pseudo panel data ought to be constructed with variables that can be assumed to be time invariant, attitudes and detailed spatial characteristics are not best-suited to this purpose (Weis & Axhausen, 2009). However, spatially differentiated studies on TB trends can be conducted. Scheiner (2010) used the national travel survey in Germany to examine TB trends between 1976 and 2002, hereby differentiating between city size categories. They found that the gap in motorisation rates between cities and suburban/rural areas has become wider, an indication that the BE affects car use. Matas & Raymond (2008) use cross section data from the Spanish Household Surveys of 1980, 1990 and 2000. They found that car ownership levels among households living in rural areas appear less sensitive to changes in income than their counterparts living in urban areas.

Even though valuable information about trends can be deduced, inferring causality with this design on the BE-TB is limited. To increase validity, multiple measurements can be conducted before and after the interventions. If a consistent break in the trend of TB occurs after the intervention, a causal effect is more likely. Furthermore, confounding variables can be controlled for to a certain extent by incorporating them in the research. When TB effects of specific BE changes are examined, one important issue is that the two independent samples should have a comparable level of exposure to the BE changes. The distance to the location of the intervention can be used as a (dose-response) proxy for the level of exposure (Ogilvie *et al*., 2006).

4 Conclusions

There is no question that BE characteristics influence TB to a certain degree. The nature and extent of causality on this BE-TB link, however, remain subject of debate. In this article we have provided a classification of multi-period designs and discussed the advantages as well as the disadvantages with regard to causality and data collection. Below, we will identify several avenues for future work and related challenges.

We demonstrate that there is a trade-off between higher ability to infer causality and lower practical applicability of research designs. We propose that an appropriate design should be selected on the basis of:

1. research aim and requirements; a causal or associative research;
2. practical constraints and opportunities: time and money constraints and existing datasets.

4.1 Opportunities for Causal Research

If the aim is to infer causality, future research should focus more on examining the impact of BE *changes* on TB. Firstly, this change could be brought about by changes in land use (e.g. new retail facilities, traffic calming schemes). In this case, we would recommend natural experimental designs with comparison groups (if possible). Secondly, people can experience changes in BE context due to residential moves or job changes. Here, we recommend prospective designs with comparison groups.
Furthermore, the aforementioned BE changes seem to make people more receptive to behavioural change interventions such as promotional campaigns. This provides an opportunity for a (largely unexplored) new research direction where changes in the BE are combined with deliberate interventions to change TB. This would enable researchers to assess the role of the BE and determine possible interaction effects with the interventions. Within this setting, we would suggest using randomised experimental designs.
Additionally, the influence of the BE can be assessed as an explanatory variable in the context of naturally occurring changes in household circumstances such as the arrival of a new baby or reaching old age. In these cases, we would recommend prospective designs with comparison groups. However, as the desired time span for the study becomes larger, the risk of drop-out will increase, and researchers may have to resort to less rigorous designs.
In all cases it is preferable to measure TB effects at (at least) three points in time, in order to distinguish between short-term and long term effects. However, even these more rigorous designs do not eliminate the possibility of third variable bias. We therefore recommend the inclusion of known TB determinants in the research (see Figure 1) and multivariate analysis to identify third variable influences and the mechanisms that mediate causality.

4.2 Practical Constraints and Opportunities

While designs with multiple observations of the same participants over time provide opportunities to infer causality, they generally also take more time, require longer term motivation / support and are more complex. This complexity relates to:

1. The increase in non-response in subsequent research waves (drop-out). Providing financial and material incentives, e-mails and phone calls to remind respondents proved to be effective in reducing drop-out rates (see Yanez, 2010 and Ortúzar, 2011).
2. The inability to select a proper comparison group/area. In this case researchers can resort to alternative approaches such as using data and information from alternative sources and including multiple baseline and follow-up measurements to identify a break in TB trends (Merom, 2003; Graham-Rowe *et al*., 2011).
3. The timing of baseline and follow-up measurements. It is often difficult to assess in advance which TB effects (or other effects, such as residential relocation, job changes, destination choices, attitude changes) can be expected over what timespan (Chen & Chen, 2009). This makes it difficult to determine the timing of follow-up measurements for the assessment of short-term and long-term effects.

Furthermore, a research question does not always require causal inferences – the desired time span may be too large and/or practical constraints may prohibit the use of rigorous designs. In these cases, retrospective and repeated cross-sectional designs can provide reasonable alternatives that still allow the study of changes over time. Researchers can resort to retrospective studies when (1) a study aims to assess changes at the level of individuals or households, (2) results need to be obtained in a relatively short time period, or (3) the BE changes being researched have already occurred. We believe this design is only reliable enough to measure more general changes such as changes in car ownership, regular mode choice for journeys to work, job (location) changes and residential moves. To apply retrospective design to assess changes in attitudes or detailed changes in TB is probably more risky. Researchers can fall back on repeated cross-sectional designs when (1) the assessment of aggregate population changes or trends is sufficient, and/or (2) only data from a continuous travel survey is available. Insights from these studies may later be used as a starting point for studies that use a more rigorous design.

This article shows that future studies to refine our understanding of causality on the link between the built environment and travel behaviour should use rigorous multi-period research designs. The classification of research designs presented in this article will aid researchers to select an appropriate design.

Acknowledgements

This article is funded by the Netherlands Organisation for Scientific Research (NWO)

REFERENCES

Aditjandra P.T., Cao J. and Mulley C. (2012) Understanding neighbourhood design impact on travel behaviour: An application of structural equations model to a British metropolitan data, *Transportation Research Part A*, 46(1), pp. 22–32.

Arentze, T., Borgers, A., Ponje, M., Starns, A. and Timmermans, H. (2001) Assessing Urban Context Induced Change in Individual Activity Travel Patterns: Case Study of New Railway Station, *Transportation Research Record*, 1752, pp. 47-52.

Bamberg, S. (2006), Is a residential relocation a good opportunity to change people's travel behaviour? Results From a Theory-Driven Intervention Study, *Environment and Behaviour*, 38(6), pp. 820-840.

Beige, S., Axhausen, K.W. (2012) Interdependencies between turning points in life and long-term mobility decisions, *Transportation*, 39(4), pp. 857–872.

Ben-Elia E. and Ettema D. (2011) Changing commuters' behavior using rewards: A study of rush-hour avoidance, *Transportation Research Part F*, 14(5), pp. 354–368.

Behrens R. and Del Mistro R. (2010) Shocking habits: Methodological issues in analysing changing personal travel behaviour over time, *Journal of Sustainable Transportation*, 4(5), pp. 253-271.

Boarnet, M.G., Anderson, C.L., Day K., McMillan, T. and Alfonzo M. (2005) Evaluation of the California Safe Routes to School legislation: urban form changes and children's active transportation to school, *American Journal of Preventive Medicine*, 28(2S2), pp. 134–140.

Boarnet M.G. (2011) A Broader Context for Land Use and Travel Behavior and a Research Agenda, *Journal of the American Planning Association*, 77 (3), pp. 197-213.

Bohnet, M., and Gertz C. (2010) Model Event History of Car and License Availability: How Accessibility Shapes Acquisition and Disposal of Cars, *Transportation Research Record*, 2156, pp. 120–130.

Bohte, W. (2010) *Residential Self-selection and travel: The relationship between travel-related attitudes, BE characteristics and TB*, Doctoral dissertation, Sustainable Urban Areas 35 (Delft: IOS Press).

Bradley, M. (1997) A practical comparison of modeling approaches for panel data. In: Golob, T.F., Kitamura, R., Long, L. (eds.), *Panels for Transportation Planning*, pp. 281–304 (Boston, Kluwer).

Cao, X., Mokhtarian, P.L., Handy, S. (2007) Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach, *Transportation*, 34(5), pp. 535–556.

Cao, X., Mokhtarian, P.L., Handy, S. (2009) Examining the impact of residential self-selection on travel behavior: a focus on empirical findings, *Transport Reviews*, 29(3), pp. 359-395.

Chatterjee K. (2011) Modelling the dynamics of bus use in a changing travel environment using panel data, *Transportation*, 38(3), pp. 487–509.

Chen, C. and Chen, J. (2009) What is responsible for the response lag of a significant change in discretionary time use: the built environment, family and social obligations, temporal constraints, or a psychological delay factor? *Transportation*, 36(1), pp. 27–46.

Chatman, D.G. (2009) Residential choice, the built environment, and nonwork travel: evidence using new data and methods, *Environment and Planning A*, 41(5), pp. 1072-1089.

Clarke, P., Ailshire, J.A. and Lantz, P. (2009) Urban built environments and trajectories of mobility disability: Findings from a national sample of community-dwelling American adults (1986–2001), *Social Science and Medicine*, 69(6), pp. 964–970.

Evenson, K.R., Herring, A.H. and Huston, S.L. (2005) Evaluating Change in Physical Activity with the Building of a Multi-Use Trail, *American Journal of Preventive Medicine*, 28(2S2), pp. 177–185.

Ewing, R. and Cervero R. (2010) Travel and the Built Environment, *Journal of the American Planning Association*, 76(3), pp. 265-294.

Festinger, L. (1957) *A Theory of Cognitive Dissonance* (Stanford University Press, Stanford, CA).

Giles-Corti B., Knuiman, M., Timperio, A., Van Niel, K., Pikora, T.J., Bull, F., Shilton, T., and Busara, M. (2008) Evaluation of the implementation of a state government community design policy aimed at increasing local walking: design issues and baseline results from RESIDE, Perth Western Australia. *Preventive Medicine*, 46(1), pp. 46-54.

Goodman, A., Sahlqvist, S., Ogilvie D. (2013) Who uses new walking and cycling infrastructure and how? Longitudinal results from the UK iConnect study, *Preventive Medicine*, 57(5), pp. 518–524.

Graham-Rowe E., Skippon, S., Gardner, B. and Abraham, C. (2011) Can we reduce car use and, if so, how? A review of available evidence, *Transportation Research Part A*, 45(5), pp. 401–418.

Handy S., Cao, X. and Mokhtarian P. (2005) Correlation or causality between the built environment and travel behavior? Evidence from Northern California, *Transportation Research Part D*, 10(6), pp. 427–444.

Kitamura, R. (1990) Panel analysis in transportation planning: an overview. *Transportation Research Part A*, 24A(6), pp. 401–415.

Krizek, K. (2003) Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? *Journal of the American Planning Association*, 69(3), pp. 265-281.

Krizek, K., Barnes G. and Thompson K. (2009) Analyzing the Effect of Bicycle Facilities on Commute Mode Share over Time. *Journal of Urban Planning and Development*, 135 (2), pp. 66-73.

Lovejoy, K., Sciara, G.-C., Salon, D., Handy, S., Mokhtarian, P. (2013) Measuring the impacts of local land-use policies on vehicle miles of travel: The case of the first big-box store in Davis, California. *The Journal of Transport and Land Use*, 6(1), pp. 25–39.

Maat, K. and Timmermans, H. (2009) Influence of the residential and work environment on car use in dual-earner households, *Transportation Research A: Policy and Practice*, 43(7), pp. 654–664.

Matas, A. and Raymond, J.L. (2008) Changes in the structure of car ownership in Spain, *Transportation Research Part A*, 42(1), pp. 187–202.

Meurs, H. (1991) *A panel data analysis of travel demand*, Unpublished Ph.D. thesis (University of Groningen, The Netherlands).

Meurs, H. and Haaijer R. (2001) Spatial structure and mobility, *Transportation Research Part D, Transport and Environment*, 6(6), pp. 429–446.

Merom, D., Bauman, A., Vita, P. and Close G. (2003) An environmental intervention to promote walking and cycling—the impact of a newly constructed Rail Trail in Western Sydney, *Preventive Medicine*, 36 (2), pp. 235–242.

Mokhtarian, P.L. and Cao, X. (2008), Examining the impacts of residential self-selection on travel behaviour: a focus on methodologies, *Transportation Research Part B*, 42(3), pp. 204-228.

Ogilvie, D., Mitchell, R., Mutrie, N., Petticrew M. and Platt S. (2006) Evaluating Health Effects of Transport Interventions: Methodologic Case Study, *American Journal of Preventive Medicine*, 31(2), pp. 118–126.

Ogilvie, D, S. Griffin, A. Jones, R. Mackett, C. Guell, J. Panter, N. Jones, S. Cohn, L. Yang and C. Chapman (2010) Commuting and health in Cambridge: a study of a 'natural experiment' in the provision of new transport infrastructure. BMC Public Health 2010, 10:703.

Ogilvie D, Bull, F., Cooper, A., Rutter, H., Adams, E., Brand, C., Ghali, K., Jones, T., Mutrie, N., Powell, J., Preston, J., Sahlqvist, S. and Song Y. (2011) Evaluating the travel, physical activity and carbon impacts of a 'natural experiment' in the provision of new walking and cycling infrastructure: methods for the core module of the iConnect study, *BMJ Open*, 2012, 2:e000694.

Ortúzar, J. de D., Armoogum, J., Madre J.-L. and Potier F. (2011) Continuous Mobility Surveys: The State of Practice, *Transport Reviews*, 31(3), pp. 293-312.

Remler D.K. and Van Ryzin G.G. (2010) *Research methods in practice: strategies for description and causation* (Thousand Oaks: Sage Publications).

Riley, D.L., Marka, A.E., Kristjansson, E., Sawadad, M.C. and Reid R. D. (2012) Neighbourhood walkability and physical activity among family members of people with heart disease who participated in a randomized controlled trial of a behavioural risk reduction intervention, *Health and Place*, 21, pp. 148–155.

Scheiner, J. (2010) Interrelations between travel mode choice and trip distance: trends in Germany 1976–2002, *Journal of Transport Geography*, 18(1), pp. 75–84.

Singleton Jr., R.A. and Straits, B.C. (2005) *Approaches to Social Research*, fourth ed (New York Oxford University Press).

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002) *Experimental and quasi-experimental designs for generalized causal inference* (Boston, MA: Houghton-Mifflin).

Shadish, W. R. (2011) Randomized Controlled Studies and Alternative Designs in Outcome Studies: Challenges and Opportunities, *Research on Social Work Practice*, 21(6), pp. 636-643.

Stopher, P., E. Clifford, N. Swann, Y. Zhang (2009) Evaluating voluntary travel behaviour change: Suggested guidelines and casestudies, *Transport Policy*, 16(6), pp. 315–324.

Transportation Research Board (2009) *Driving and the Built Environment: The Effects of Compact Development on Motorized Travel, Energy Use, and CO2 Emissions*. Special Report 298 (Washington, DC: Transportation Research Board of the National Academies).

Van Wee, B. (2011) Evaluating the impact of land use on travel behaviour: the environment versus accessibility, *Journal of Transport Geography*, 19(6), pp. 1530–1533.

Verhoeven, M. (2010) *Modelling life trajectories and transport mode choice using Bayesian Belief Networks*. Dissertation faculty of architecture building and planning, Bouwstenen nr 143 (Eindhoven: University of Technology Press Facilities).

Weis, C. and Axhausen, K.W. (2009) Induced travel demand: Evidence from a pseudo panel data based structural equations model, *Research in Transportation Economics*, 25(1), pp. 8–18.

Woldeamanuel, M.G., Cyganski, R., Schulz, A. and Justen, A. (2009) Variation of households' car ownership across time: application of a panel data model, *Transportation*, 36(4), pp. 371–387.

Yanez, M. F., Mansilla, P. and Ortúzar, J. de D. (2010) The Santiago Panel: measuring the effects of implementing Transantiago, *Transportation*, 37(1), pp. 125–149.