Watching passengers:
Using structured observation methods on public transport

Ms Marie Russell. PhD Student. Visiting student at the Centre for Transport & Society, University of the West of England. Registered for PhD at University of Otago, Wellington, New Zealand.

Abstract
Direct observation of human behaviour is used in many disciplines. Ethnographic observation, a qualitative method, has been used successfully in public transport passenger studies. Structured observation is an unobtrusive method which can yield large amounts of quantitative data. This paper discusses structured observation as a research method with public transport passengers.

Designing a structured observation study requires decisions about (1) sampling: what kinds of passengers to observe, what modes and routes, times of day, week, and season to conduct observations; (2) categories: what behaviours and factors to observe; (3) data collection: how will behaviours and other data be recorded; and (4) analysis and reporting.

The paper compares and critiques four recent studies which used structured observations to examine passengers' travel time activities: by Timmermans and Van der Waerden (2008); Ohmori and Harata (2008); Thomas (2009); and the author of this paper (Russell, In Press).

One of the areas of practical difficulty is data collection. Observing seated passengers on a long train or ferry journey may be relatively easy, but the crowding and intense activity on a rush-hour bus make methodical data collection very challenging.

A structured observation study of public transport passengers can only answer questions about manifest behaviour; it tells us nothing about motivations, feelings or attitudes. Those topics can only be deeply understood through non-observational research methods. But as well as providing extensive data, the experience of structured observation field work gives an excellent background to researchers wishing to understand the everyday experiences of passengers.

Background

It is common in transport appraisals to focus on travel time savings as a key benefit of transport investments, thereby treating travel time as essentially wasted time (Hensher, 2001a, Hensher, 2001b, Mackie et al., 2001, Metz, 2008, Mokhtarian, 2005, Wardman, 1998, Wardman, 2001). There is, however, very little research to justify this valuation. My PhD research in Public Health investigates how public transport passengers in two New Zealand regions use and value their travel time, and its impact on health and wellbeing; a social, not an economic study. As a preliminary investigation, I undertook structured observations of bus and train passengers around Wellington, New Zealand, during November-December 2008, to examine how passengers were spending their travel time. This paper discusses the use of passenger observation in transport research in general, describes the structured observation method, compares the different approaches discussed in the literature and reflects upon some of the issues encountered using this method.

Ways of observing passengers

There are few discussions of observation as a method in the literature about passengers' activities. Ten years ago Clifton and Handy (2001) wrote that participant observation, although used elsewhere in urban studies, 'has not often been used in travel behaviour research' (Clifton and Handy, 2001:10). A prime consideration is that observation is not appropriate if we want to find out what passengers are thinking or feeling; it can only be used to assess manifest behaviour. And, usually, we simply cannot interpret observed behaviour: for example, a person reading a novel on the train could be doing so for leisure, or for study, or even for work.

Bryman (2008:257) identified major types of observation research used in social research generally; there is some cross-over between these: Participant observation; Non-participant

observation; Unstructured observation; Simple observation and contrived observation, and Structured or systematic observation.

Participant observation is associated with qualitative methods where the researcher is immersed over time in a setting or group, often as an 'insider', to elicit meanings and understand nuances of behaviour, ideas or emotions. Anyone coming from a background in anthropology or sociology, cultural studies or similar fields may be familiar with participant observation and ethnographic observation[1] made famous in studies like Margaret Mead's 'Coming of age in Samoa'. Those with a background in sciences like biology or botany may be familiar with structured observation. A researcher counting and recording the size of beetles in a square metre of grassland is conducting a form of structured observation.

In the social sciences, naturalistic observation is assumed to 'not interfere with the people or activities under observation' (Angrosino, 2005:730) and people 'are free to vary their individual and social responses' (Sackett et al., 1978:2). Knowing they are being observed can and does affect how people behave (Patton, 2002:269). To systematically observe passengers in a covert way, a hidden video camera might be used. But there are methodological, cost and ethical reasons against this approach (Sackett, 1978).

Ethnographic observation studies on public transport

Several ethnographic observational studies of public transport passengers and travel have been carried out (e.g. Delannay, 2001, Fink, 2006, Nash, 1975, Watts, 2008, Jain, 2009). Ethnographic research has its own requirements, protocols and rigour. It involves detailed observation, and focuses on nuances of behaviour in a specific setting.

Nash's work on riding buses in Tulsa, USA had the intriguing title: 'Bus riding: community on wheels' (1975). He saw bus-riding as 'a peculiar urban form of belongingness' (Nash 1975:100) and explored the 'taken-for granted knowledge and membership feelings of participants within a particular urban setting' (Nash 1975:122). Nash's lovingly-reported observations were made over a two-year period riding various routes. He observed different behaviours and relationships among old-timer and new drivers, 'regulars' i.e. passengers who have 'negotiated a routine of riding' (Nash 1975:101) and 'newcomers' (who have not); and other behaviours such as working out the timetable and timing the bus; hailing a bus; and choosing a seat. Nash described 'rider membership' of that community on wheels; and the circumstances that work against community: 'New lines, crowded buses, new riders, and new drivers' (Nash 1975:118).

Delannay in her unpublished work on 'the social organization of riding the bus' (2001) emphasised the care passengers took to select a seat so as to maintain anonymity in the bus as 'a world of strangers' in Lofland's phrase; a 'cocoon of privacy' achieved 'by carrying out civil inattention' to others. The bus was 'not some asocial abyss', but had social and personal meanings. Similarly to Jain & Lyons (2008), Delannay, from her observations of riding buses in Madison, Wisconsin USA, saw bus travel time as a 'transition space: a space without meaning or purpose for its own sake' (Delannay, 2001:30). The transition space was one of daydream, speculative experience, 'relief from [passengers'] work-worlds or home-worlds', and anonymity. Another study of big city buses was conducted by Fink in Los Angeles where she was a reluctant bus-rider. She used 'ethnographic fieldwork methods to explore and understand bus culture and behavior' (Fink, 2006: Abstract).

In Britain, Jain and her colleagues saw buses as neglected in mobilities studies. They travelled on buses and made use of 'mobile ethnography' (Jain, 2009). Watts' ethnography of train travel was 'a travelogue of one train journey across England' (Lancaster to Penzance). She used 'ethnographic evidence in the form of quotations and photographs woven through the text' (Watts, 2008:711). Watts mentioned the

[1] Ethnography: *a research method in which the researcher immerses him or herself in a social setting for an extended period of time, observing behaviour, listening to what is said in conversations between others and with the fieldworker, and asking questions.... 'an ethnography' is frequently used to refer to the written output of ethnographic research*. BRYMAN, A. 2008. *Social research methods,* Oxford, Oxford University Press. p693.

need to negotiate with the train company before the journey. Her interesting reflection on the researcher's role noted some of the challenges of a researcher working solo:

> *As an ethnographer I was neither a passenger nor a member of the train crew, my labours, what I did, were different and marked me out. I was not immobile like an unpacked passenger, nor was I responsible for the mobility of the train.... My work onboard created a different temporality to those around me... I had to continually work at locating myself as an ethnographer and resist shifting to the location of a passenger* (Watts, 2008:718).

Ethnographic accounts provide fascinating insights into passenger behaviour and experiences. But ethnography and related types of deeply qualitative observation cannot yield information about the *range of activities* among *large numbers* of bus and train passengers, or show which behaviours are more common, or how they are shared across different population groups and different modes.

Watts and Urry (2008) discussed methods for studying travel time use: large surveys of passengers, focus groups, key informant interviews, ethnography, travel diaries and short videos. A method they did not use and which I now discuss here is 'structured observation'.

Structured observation

Structured observation, often used in schools (Bryman, 2008:270) has been used in studies of travel time use, and other areas of transport research, for example, a study assessing public transport passenger safety behaviour and risk (Daniel et al., 2009).

For my PhD research I have used a sequential mixed methods design, with three phases of data collection, each using a different method: the first phase comprised observations of 812 passengers, the second comprised qualitative interviews with 48 passengers, and the last phase was a pen-and-paper self-completion survey of over a thousand passengers.

Bryman describes structured observation as a technique where researchers use

> *explicitly formulated rules for the observation and recording of behaviour. The rules inform observers about what they should look for and how they should record behaviour. Each….*'participant'… *is observed for a predetermined period of time using the same rules. These rules are articulated in what is usually referred to as an observation schedule.... [whose] aim... is to ensure that each participant's behaviour is systematically recorded so that it is possible to aggregate the behaviour of all those in the sample in respect of each type of behaviour being recorded.* (Bryman, 2008:257).

Why use structured observation with passengers?

Anyone who pays the fare can freely observe passengers' behaviour on the bus or train. So, why would one want to use structured observation as a method in public transport research? Key reasons for employing such methods include that: it is an unobtrusive method (takes place in public; there are few ethical issues, e.g. no need for personal consents unless using video/ photographic/ audio recording); it takes place in a 'natural setting'; passengers generally do not know they are being observed and they behave naturally; it can yield a large amount of robust quantitative data in a relatively short time; it provides an excellent familiarisation with public transport by getting the researcher out into the field and examining behaviours and environments in a methodical way; it avoids the subjectivity and recall bias of data obtained through survey and interview methods and travel diaries; It is fairly economical, and it can provide snapshots of behaviour or it can provide longitudinal data about passengers on a trip or series of trips.

But the disadvantages of structured observation also need to be considered: it only provides information on passengers' manifest behaviour, i.e. it tells us nothing about people's subjective experience, their feelings, attitudes, reasons, wishes and preferences, fears, desires, opinions, etc; there can be some challenging issues in data collection: e.g. the researcher has no control over the passengers. Like the beetle flying away as the biologist tries to see how big it is, passengers get on and off the bus or train according to their needs, not the researcher's. Further, observers' reliability and consistency may vary over time; two

observers may disagree. These issues can be mitigated but need to be addressed. Overall, a structured observation study may be done fairly quickly. The expensive parts of the research are the labour-intensive parts: data collection, data entry and analysis.

Key steps in a structured observation study

Decisions about the design come back, as in all research, to one's research questions, and a time-honoured test is to ask 'what do I want to know, and why?' (Jenny Neale). In observation studies, the approach, categories, sampling frames, timeframes etc. all depend on whether one is interested in breadth across a population or depth; in specific behaviours or an open-ended range of behaviours; in behaviour over time or in a snapshot. Designing a structured observation study of passengers requires decisions about (1) sampling; (2) categories; (3) data collection; and (4) analysis and reporting. Some of the issues are discussed in more detail below, with examples from several studies.

Sampling and sampling frames in recent observation studies of passenger activities

It is important to establish some kind of *sampling frame*: a 'listing of all the units in the population from which the sample will be selected' (Bryman, 2008:168). The sampling frame and sample indicate *where* and *when* to observe. In a public transport context this could include identifying modes, routes, days of the week, times of day, seasons, or types of passenger. Quota sampling might be used here; meaning, for example, observing only school pupils in a certain city, or people on train routes to wealthy and poor areas. Thinking through the rationale for items in the sampling frame entails linking them to the overall research question, justifying the selection, and stating clearly how and why they have been chosen.

Bryman notes that 'Studies in public areas...do not permit random sampling' (Bryman, 2008:263). This kind of research with public transport passengers generally uses non-probability or strategic sampling. Some of the types of sampling listed by Martin and Bateson (1986:Chapter 4) are: *Ad libitum* sampling (whatever's happening at the time); Focal sampling (specific individual observed for a period of time); Scan sampling (group scanned at set intervals and behaviour recorded); and Behaviour sampling (entire group watched and a particular behaviour recorded).

Examples of sampling frames: *where* and *when* are we observing?

Timmermans & Van der Waerden studied the 'extent and nature of activity engagement while travelling on short-distance trains', including the effect of demographic and contextual variables on the activities (2008:2). They observed passengers on one day in June 2007 on a particular line of the Bay Area Rapid Transport System San Francisco, in the early morning peak, middle of the day and late afternoon, each time on a carriage 'selected at random'.

Ohmori & Harata's aim was similar: they wanted to see what activities commuters were engaging in on the train, and what factors of the travel environment and individuals caused the difference (2008:548). Before conducting a survey, they observed passengers from 6.30am to 7.04am on 11 weekdays in November and December 2003, on a specific train route, the Odakyu Support No 60 train from Machida Station to Shinjuku Station on the Odakyu-Odawara line (a high-grade train where for an extra fee people are guaranteed a seat). They observed six to eight people every one minute during the journey.

Thomas' interest for his study in psychology was in passengers' adaptation to interpersonal discomfort on public transport in terms of *'defensive and social adaptations'* (Thomas, 2009:37). He observed passengers in Wellington, New Zealand over eight winter days from 6.30am to 6.00pm on 38 bus trips and 23 train trips. A random selection was made of bus route numbers and all four train lines to Wellington were included (Thomas, 2009:38). Peak (7.00am-9.00am and 3.00pm-6.00pm) and off-peak times were included. Thomas observed all passengers who entered a bus or half of a train carriage.

The aim of my research was to assess the frequency of passenger activities during bus and train travel, using structured observations of passengers in a purposive sample of bus and train routes and times, in the Wellington area. I observed passengers over nine days in November-December 2008 on 24 bus and 22 train trips, both short (20 minute) and long trips

(up to two-hours), in downtown and suburban routes, encompassing wealthier and poorer areas, according to the NZ Index of Deprivation (Salmond et al., 2007). Morning (before 9.00am) and evening (3.00pm-6.30 pm) peak commuting times were included (New Zealand Transport Agency, 2008), but also several night and middle-of-the-day times. Two observers recorded two different passengers each five times over four minutes, on the minute.

Categories in recent observation studies of PT passenger activities

The observation schedule provides the categories of *who* and *what* to observe. This can include information that provides context. For example, if we think they affect our topic, we might record factors like the weather or the degree of crowding on a vehicle. The type of person being observed is important if we want to know what kinds of people use a computer on the train, say, or if men are more likely than women to listen on headphones.

Examples: *Who* are we observing*?*

Timmermans & Van der Waerden recorded for each person observed: gender; race (Caucasian or Other); age (under 18, 18-25, older than 25); travel party (alone, couple, other); trip duration (less than 20 minutes, 20-30 minutes, more than 30 minutes); time of day.

Ohmori and Harata's report suggests they did not observe passenger characteristics, but they followed their observation study with a survey (June 2004) where they *did* ask for demographic information: age, sex, occupation, income, work hours, overtime allowance and many other aspects.

Thomas recorded characteristics of the person observed (age, gender, part of a couple), and other aspects: where they were in the vehicle (standing, sitting, seat location, seating choice: next to same gender person, next to person using defensive behaviours, seat-shifting), and interpersonal distance.

In my study only adults were observed. We noted gender and broad age group (young; i.e. about 18 to 30-35; middle aged: about 35 to 60; and older: over 60). In New Zealand it is considered inappropriate to guess at people's ethnicity; it is constructed as meaningful only through self-identification (Statistics New Zealand, 2005) so race/ethnicity were not included.

Examples: *What* are we observing*?*

In passenger activities research, the coding schedule of *what* to observe is crucial. Unless all behaviours of a person or group are being recorded (as in *ad libitum* or focal sampling), a schedule will list the behaviours or activities to record. Interestingly, in the studies I reviewed, there was considerable accord between the categories of train passenger activities used in studies in Japan (observation and survey) (Ohmori and Harata, 2008); the USA (Timmermans and Van der Waerden, 2008), and New Zealand (Thomas, 2009). Two other studies: surveys, *not* observational studies, in Great Britain (Lyons et al., 2007) and Norway (Gripsrud and Hjorthol, 2009), show similar categories.

Table 1 shows the categories used in a selection of travel time activity studies: most were observational studies. There were some subtle differences, e.g. the activity called 'window gazing /people watching' in the UK survey (Lyons et al., 2007) is 'Seeing advertisements, scenery and people' for Ohmori and Harata (2008). In New Zealand, Thomas did not record this as his interest was in defensive behaviours. I called it 'Looking ahead/ out window'.

Thomas also noted the weather (observed weather and official measured weather), verbal interaction, defensive behaviours (sitting on aisle seat, bag on seat or enlarging posture), and included activities under the heading of 'situational withdrawal into an activity' (the activities were listening to music, reading, phone texting, phone talking, head down likely to be reading, sleeping, knitting, laptop use).

Table 1. Selected activity categories; studies in Japan, USA, UK, Norway, New Zealand

Passenger activity category	Japan	USA	NZ (Thomas)	NZ (Russell)	UK	Norway
Reading newspaper, book etc.	X	X	X	X	X	X
Talking to other passengers	X	X	X	X	X	X
Sleeping/ snoozing	X	X	X	X	X	X
Listening to music/ radio	X	X	X	X	X	X

Window gazing/ looking ahead	X	X		X	X	X
Working/ studying					X	X
Talking on Phone	X	X	X	X	X	X
Text messaging	X	X	X	X	X	X
Personal care		X				
Work computer		X	X	X		
Game (various)		X			X	
Romancing		X				
Eating/ drinking	X			X	X	
Smoking cigarettes	X					
Singing songs	X					
Thinking	X				X	X
Using PC/ PDA, video, game	X				X	X
Care of children					X	X
Knitting, needlework			X			X
Writing			X	X		
Handling wallet, etc				X		
Other (describe)				X		X

In their pilot list, Timmermans and Van der Waerden included several items which I would suggest should not be used, for example: *Doing nothing; Travelling only; Reading work material; Phoning– professionally; Phoning– socially; Phoning– reason unknown; Being called– professionally; Being called– socially; Being called– reason unknown.* A person sitting looking ahead is not necessarily doing nothing; and one cannot judge from observation alone the reason for a phone call. Timmermans and Van der Waerden wisely condensed such categories at analysis.

Note that the categories may reflect different cultural practices (the Japanese study includes 'singing'; as an activity) and varying national regulatory differences (for example, about smoking, which is banned on public transport in some countries).

In New Zealand, after initial casual observation and looking to the literature, notably the UK surveys, I developed an activities list which was then tested. We added a category 'handling wallet, equipment, etc.' after our pilot study, having observed people rummaging in their bag, wallet or purse. We added 'other (describe)', which was needed and used several times.

Data collection

Some thought needs to be given before going into the field to ethical issues and transport company consents. Watts discussed the need to contact transport companies before starting her ethnography. In Wellington, I contacted the relevant public transport providers to explain the research. The managers of both bus and train operations generously offered free passes for the two researchers, and a letter of consent and support. Researcher safety should be considered: in my case, two women researchers worked together for safety reasons, avoiding late night trips. Observers may need to answer passengers' questions: 'Are you doing a survey?' 'Can you just make the trains run on time?' and queries from bus drivers and train conductors. It is worthwhile maintaining friendly and polite relations with inquirers, for the smooth running of the project, but also because we have obligations as part of the research community. If we or our assistants are in any way rude or unfriendly, we will surely queer the pitch for future researchers.

To record the data, some sort of recording template is needed, in paper or in electronic form. Thomas used a diagram representing the normal seating plan in buses and trains, with room to write in information. In my research where Summer Student, Rachel Price and I were observing two people each at a time, we used a clip board with paper sheets, each of which had room for the demographic details of two people and associated activity checklists.

Piloting the template and the researchers in the field is crucial. It can reveal unforeseen problems with the template or the research design. Timmermans and Van der Waerden, as noted above, evidently realised *after* the fieldwork that you cannot tell if a phone call is social or for work, and this may partly explain why they refer to their study as a pilot.

The pilot in my research was extremely important as it showed that even after repeated attempts over a week, we could not make the initial data collection plan work. It was

originally intended that two researchers sit or stand together on the public transport vehicle, then, at an agreed time, and beginning with the same passenger, each observe and record (using pen and paper) all the passengers in the vehicle, one after the other. For each passenger, their general age range and their gender would be noted, and whether or not they appeared to be a 'single' or a 'with' (Goffman's 1971 terminology for people who are in couples or a group was 'withs', as opposed to 'singles' [Nash 1975]), and what they were doing. This is the general method of Timmermans and Van der Waerden (2008), and appears similar to that of Thomas (2009). But it just did not work for us, especially on buses, where passengers got on and off every few minutes and our note-taking could not keep up. One observer worked faster than the other; there was a marked lack of agreement between the observations, especially about the age of passengers (possibly owing to a 32-year age gap between the observers). We found that where vehicles were crowded in peak times, we simply could not *see* all the passengers in the vehicle. So I went back to the drawing board, redesigned the data collection protocol, and trialled it again (Russell et al., in press).

In the end, we elected to use a sampling approach where a group or a selection of people is scanned at set intervals and behaviour is recorded. Martin and Bateson (1986:54) give the various names for this approach to recording behaviour: 'instantaneous sampling', 'point sampling' or 'fixed interval time point sampling', and they also advise on choosing the sample interval. We observed individuals five times over an observation period of four minutes, first noting passenger characteristics and then, once a minute and on the minute, viewing the passenger and immediately recording the passenger activity at that instant. Four minutes for observing each passenger allowed us to record some of the variability in behaviour. It was long enough to obtain a large amount of data, but not so long that many passengers were lost to observation especially on buses.

The two observers usually started with those nearest to them on a particular side of the vehicle, taking a side each by agreement; but also bore in mind a wish to observe roughly equal numbers of men and women, and sometimes individuals were purposefully selected on the basis of gender (numbers were tallied at the end of each trip). Note that the two observers did not observe the same passengers: there was no inter-rater check. There were still difficulties, as for example when passengers boarded and stood in the aisle at peak times, completely blocking the researchers' view of passengers already under observation. One of the observers, Price, noticed that even if the observer could not see the passenger directly, bus and train windows had reflecting glass which especially at night was useful in revealing adequately what passengers were doing.

Thomas apparently observed all the passengers who boarded the vehicle (Thomas, 2009), though he does not describe his data collection method in full; he appears to have used a form of 'behaviour sampling' as described by Martin and Bateson (1986). Ohmori and Harata's observer recorded six to eight passengers' activities every minute on the minute (Ohmori and Harata, 2008). We found observing two passengers at a time on the minute was comfortable. More than two passengers at a time would be feasible in our view, or two every 30 seconds, but we believe eight per minute would be demanding. Our two-passengers, four-minutes, five-observations protocol was appropriate to elicit a large amount of data within the time and research resources available.

Problems with structured observation as a method may arise during the data collection when there is more than one observer recording the same thing, concerning the degree of agreement between them (inter-rater reliability). For rigour, having two researchers observe and record the same people and behaviour is desirable and the degree of agreement can be tested using Cohen's kappa. If there is an unacceptable lack of agreement during a pilot, additional training or calibration of the method and tools can be undertaken.

Thomas, a sole observer, had someone work with him on a small subset of his observations and then assessed the degree of inter-rater agreement in that subset using Cohen's kappa. He decided his observations were pretty accurate. Ohmori and Harata and Timmermans and Van der Waerden do not mention this issue, though the latter stated that 'some mistakes will be made' in field observations (Timmermans and Van der Waerden, 2008:3).

Other issues around data collection which may affect reliability in the research are that an observer's attention may flag (affecting intra-rater reliability), the consistency of observations over time by each observer may change (Martin and Bateson, 2007:Chapter 7), or some data missed. Thus 'observer drift' (Robson, 1993:224), 'observer fatigue' (Martin and Bateson, 2007:80) or 'observer decay' (Hollenbeck, 1978:84) may be of concern. Watts (2008) described the challenge of maintaining her ethnographer's role and location as a *researcher*, and not falling into the *traveller* role. Concerning observer fatigue, we found we could work a four or five hour period before becoming tired, but recommend a half-hour break after every two hours: this is not something you want to do all day!

Dealing with the data

Martin and Bateson (1986:Chapter 9) advise presenting results of both exploratory and confirmatory analyses (descriptive and inferential statistics). Ohmori and Harata gave descriptive statistics only and Timmermans and Van der Waerden gave frequencies as well as a stepwise discrete choice analysis – a logit analysis

> *with the probability of engagement in activity classes as the dependent variable and with gender, race, time of day and total duration of the trip as explanatory variables* (Timmermans and Van der Waerden, 2008:5),

and other cluster analyses. We entered data in an Excel spreadsheet and, like Thomas, analysed it in SPSS. Bus and train data were amalgamated to produce a single dataset, and our five time-intervals were listed in a single column for analytical purposes. We produced the descriptive statistics and then conducted binary logistic regression analysis for the association of observed activity against the demographic and travel covariates that we were interested in (gender, approximate age group, transport mode, and peak/ off-peak travel time). Odds ratios from the logistic regression were reported to examine the relationship between the covariates and each activity. A critical P-value of .05 and 95% confidence intervals were included to test for significance.

One of the things analysis needed to deal with was multitasking – people who were doing more than one activity at a time, as well as travelling. A passenger may have headphones in and be listening to music while reading a book or doing a crossword puzzle. They may be eating and texting at the same time, perhaps with their headphones in. In addition, a passenger might undertake several different activities sequentially over the observation period, for example in our study, reading at Times 1 and 2, talking at Time 3 and texting at Times 4 and 5. Or a passenger might have alighted after two minutes.

To accommodate this diversity, our data analysis referred to the numbers of passengers who were *ever observed* doing the activity. A passenger reported as 'ever-texting' may have been reading at four of the times she was observed and texting only at the fifth; or eating while texting.

An effect of the 'ever observed' approach may be to inflate some of the data. For example, in virtually every journey, a passenger is likely to look ahead or out the window at some point, and our method may count this activity more than its duration in reality. Results around this could therefore be an artefact of the method. Another category where a behaviour is so integral a part of the journey that it may be distorted in the study is the handling of wallet or purse. This is especially the case where passengers have a ticket clipped or pay cash in exchange for a paper ticket; thus handling their wallet or purse, removing money, or stowing a clipped ticket. Many passengers on both buses and trains in Wellington, however, show a pre-paid token and do not present cash or require change.

Ohmori and Harata dealt with the multitasking of passengers in a different way, through deciding which was the 'main' activity:

> *...where more than two types of activities were simultaneously conducted for each one-minute, the main activity was decided based on the following priority: (i) reading newspapers, reading magazines, reading books, reading documents, using PDA, emailing by mobile phone, web-browsing by mobile phone, sleeping, (ii) eating, smoking cigarettes, (iii) drinking, (iv) listening to music, listening to radio* (Ohmori and Harata, 2008:Table 2).

There may be a certain logic in this, but such a classification can really only be found through questioning multitasking passengers.

So how *did* people spend their travel time?

Given the differences in aims, data collection methods and analyses, it is difficult to compare results from the passenger activity studies discussed here. Further, cultural and other differences in the studies from the USA, Japan and New Zealand are significant and may render the comparison of results unhelpful. With this caveat, that the information needs to be understood as very general, key results from the studies are summarised in Table 2 below. Some results sum to 100%, others do not because of the multitasking or main/ other activites aspect of the study designs. Readers are referred to the individual reports, which can be seen online or elsewhere, except for Russell, Price et al. (in press).

Table 2. Travel time activities in observation and survey studies from five countries

	Ohmori & Harata (2008) Japan	Timmermans & Van der Waerden (2008) USA	Thomas (2009) NZ	Russell et al.(in press) NZ	Lyons et al., (2007) UK	Gripsrud and Hjorthol (2009) Norway
Number of people observed / modes	84 train	161 train	1703: Bus=1142 Train=561 434 did 'activities'	812: Bus=353 Train=459	26,221 train SURVEY	1196 train SURVEY Commuter - business
Window gazing/ looking ahead etc	16.7%	63.6%	-	65.3%	18%	50%-60%
Reading (/writing etc	61.9%	4.8%	12.7%	21.7%	34%	35%-27%
Sleeping/ snoozing	66.7%	6.1%	0.7%	8.9%	3%	36%-31%
Talking	-	13.8	23.9%	15.4%	5%	19%-35%
Listening to music/ radio	8.4%	1.6%	8.9%	19.2%	3%	26%-7%
Phone text	13.1%	2.8%	2.4%	9.2%	2%	-

Thomas conducted observations of 1703 passengers on Wellington buses and trains. Thomas was not examining the range of behaviours *per se* but looked at passenger characteristics, seat selection, movement within the vehicle, verbal interaction, and 'defensive behaviours' in which category he included listening to music, reading etc. (Thomas, 2009:42). This study by Thomas is of considerable interest to me. Thomas has not fully explained his method but it included, for most of his observations, one person observing all the passengers boarding a bus or one half of a train carriage, noting any subsequent seat changes and departure, gender, age, couple relations, seat location(s) and patterns, as well as activities such as verbal interaction, bag placement, and travel time activities. Without greater detail than is given in his thesis, it is difficult to know exactly how this was accomplished but since Thomas observed 1142 bus passengers on 38 trips, this gives an average of 30 people observed per trip; on trains the average would be 24 people per trip. Our results for activities differ from Thomas' for basically the same population. It is unclear whether these differences relate to different times of year (we collected data in summer; Thomas in winter), different times of day, or, more likely, methodological differences.

My research found that nearly two-thirds of the passengers observed spent some of their travel time looking ahead or out the window (65.3%) but this was seen more on the bus (76.5% of bus passengers) than on the train: 56.6% of train passengers were looking ahead or out at some point during the observation. Overall 21.7% of the passengers were observed reading, with more than twice the proportion seen reading on the train (28.8%) than on the bus (12.5%). Passengers ever-observed with headphones on were 20.9% of train passengers and 17% of bus passengers. Slightly more people were observed talking to other passengers on the train (16.8%) than on the bus (13.6%). Texting was more commonly observed (9.2% of all passengers) than talking on a cell-phone (1.5%). Activities observed more frequently on trains than on buses were reading, using a computer, sleeping/ eyes closed, writing and handling wallet, bag etc.

Our analysis identified how activities interacted with the demographic and contextual factors of gender, age, mode and time of day. Women were significantly more likely to be talking, and less likely to be using a computer than men. Older people were significantly less likely to be texting, using headphones, eating/ drinking, or looking ahead/ out window than younger people, but significantly more likely to be reading. As noted above, more people were looking

ahead/ out window on buses than on trains and the odds ratio for this showed a statistically significant difference. Train passengers were significantly more likely than bus passengers to be reading, using a computer, sleeping/ eyes closed, writing and handling their wallet or belongings. Time of day revealed fewer clear-cut differences, with passengers significantly more likely to use a computer at peak travel times, and more likely to be looking ahead/ out window at off-peak times of day.

Timmermans and Van der Waerden (2008) in their San Francisco study found almost all the passengers observed were 'doing nothing'. The authors reported differences in activities: 'doing nothing, sleeping, talking, reading and [listening to] music' by sociodemographic and contextual variables: gender, race, age, travel party (alone, couple or group), trip duration and time of day. That almost all of the people observed were 'doing nothing'

> *casts doubt on the prevalence of multitasking while travelling on trains, at least for this sample, which concerned travelling for relatively short distances* (Timmermans and Van der Waerden, 2008:5).

Other activities discussed were sleeping (commoner among women and non-Caucasians and in the morning commute, and less common among the 18-25 year olds (almost half of the sample); and talking (commoner among women and Caucasians).

The Japanese study by Ohmori and Harata (2008) included an observation of 84, and a survey of 503 passengers on 'normal' and 'high grade' trains. The observations showed sleeping and reading as the most frequent activities; sleeping at a high rate (67%). But the observation study did not appear to include a 'doing nothing' category. The ensuing survey evidently did have such a category, however, and a quarter to a third of passengers reported 'thinking of something' for work or leisure. Some activities differed by trip length: the longer the trip, the more likely passengers were to be sleeping or reading, especially if they had a seat. Not having a seat did not prevent sleeping, though.

Two surveys of train passengers are of interest as well in Table 2. In the large British survey (N=26,221 train passengers) about different activities while travelling, reading for leisure (34%), window gazing/ people watching (18%), and working/ studying (13%) were the frequent activities reported by passengers (Watts and Urry, 2008). For British passengers, unlike those in the USA observational study, sleeping/ snoozing happened more on the 'return' journey (Lyons and Chatterjee, 2008:192). Window-gazing was high on short journeys (Lyons et al., 2007:112) and the authors suggest there may be

> *a possible travel duration threshold below which there is not a suitable amount of time to do* other than *window gaze/ people watch* (Lyons et al., 2007:112) (their emphasis).

In Norway, Gripsrud and Hjorthol's (2009) train survey (N=1196) found well over a third of passengers using travel time for work, with nearly a quarter of commuters having their travel time paid as work time.

A neglected area of study is waiting time. A group of fourth-year medical students at the University of Otago, Wellington, carried out a structured observation study of passenger activities at bus stops and train stations in Wellington in 2009 (Elkin et al., 2009).

In summary

Structured observation is a 'way of quantifying behaviour' (Robson, 1993:206) as it 'focuses on the frequency of… actions' (Gray, 2004) and 'employs explicitly formulated rules for the observation and recording of behaviour' (Bryman, 2008:257). Unlike ethnographic studies, it produces quantitative data. At the time of my research, we had not reviewed the other studies discussed here. There is general advice about using the method, however, in particular in Martin and Bateson's books (1986, 2007). Experience, and a review of other research using structured observation on public transport, showed that care is needed in identifying the sampling frame, developing the coding schedule and testing the data collection protocols.

This kind of research seems to work well in a larger mixed methods research design. Thomas (2009) followed his observations with further surveys of passengers and non-public transport users. Ohmori and Harata (2008) also followed their observations with a self-report

questionnaire to passengers. My observations were followed by telephone interviews with passengers and then a self-report questionnaire. I found the experiences of designing the research, revising the data collection approach and actually getting out in the field in a systematic way afforded an excellent introduction to the rest of the project.

Acknowledgements*:* My observations research was supported by: Supervisors Louise Signal, Jacqueline Cumming and James Stanley; Summer Student Rachel Price and Research Assistant Zac Gerring.

References:

ANGROSINO, M. V. 2005. Recontextualizing observation: ethnography, pedagogy, and the prospects for a progressive political agenda. *In:* DENZIN, N. K. & LINCOLN, Y. S. (eds.) *The SAGE handbook of qualitative research.* 3rd ed. Thousand Oaks: SAGE Publications.

BRYMAN, A. 2008. *Social research methods,* Oxford, Oxford University Press.

CLIFTON, K. J. & HANDY, S. L. 2001. Qualitative methods in travel behaviour research. *International Conference on Transport Survey Quality and Innovation.* Kruger National Park, South Africa.

DANIEL, J. R., ROTTER, N. G. & NICHNADOWICZ, V. F. 2009. Customer behavior relative to gap between platform and train. New Jersey Department of Transportation Division of Research and Technology and U.S. Department of Transportation Federal Highway Administration.

DELANNAY, M. C. 2001. Maintaining anonymity: the social organization of riding the bus (unpublished thesis).

ELKIN, L., MCLACHLAN, A., CLELAND, A., HENARE, E., SINGH, V., BINNS, H., KAUTOKE, S., PARSONS, E., RAMSDEN, T., WILSON, A., NARAN, J., ALY, S. S. A. & CHEN, L. 2009. Watching and Waiting: An Observational Study. Wellington: University of Otago, Wellington.

FINK, C. N. Y. 2006. Presentation of self in everyday transit life: ethnographic study of Los Angeles bus culture *TRB 85th Annual Meeting Compendium of Papers.* Washington DC: Transportation Research Board

GRAY, D. E. 2004. *Doing research in the real world,* London, SAGE Publications.

GRIPSRUD, M. & HJORTHOL, R. 2009. Working on the train: from 'dead time' to contractual time. *Network- ICT: mobilizing persons, places and spaces. Fourth specialist meeting of the Network* Quebec Institute of Transport Economics.

HENSHER, D. A. 2001a. Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy,* 35**,** 71-98.

HENSHER, D. A. (ed.) 2001b. *Travel behaviour research: the leading edge,* Amsterdam: Pergamon.

HOLLENBECK, A. R. 1978. Problems of reliability in observational research. *In:* SACKETT, G. P. (ed.) *Observing behavior: v 2 Data collection and analysis methods.* Baltimore: University Park Press.

JAIN, J. 2009. The making of mundane bus journeys. *In:* VANNINI, P. (ed.) *The cultures of alternative mobilities: routes less travelled.* Farnham: Ashgate.

JAIN, J. & LYONS, G. 2008. The gift of travel time. *Journal of Transport Geography,* 16**,** 81-89.

LYONS, G. & CHATTERJEE, K. 2008. A Human Perspective on the Daily Commute: Costs, Benefits and Trade-offs. *Transport Reviews,* 28**,** 181 - 198.

LYONS, G., JAIN, J. & HOLLEY, D. 2007. The use of travel time by rail passengers in Great Britain. *Transportation Research Part A,* 41**,** 107-120.

MACKIE, P. J., JARA-DÍAZ, S. & FOWKES, A. S. 2001. The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review,* 37**,** 91-106.

MARTIN, P. & BATESON, P. 1986. *Measuring behaviour: an introductory guide,* Cambridge, Cambridge University Press.

MARTIN, P. & BATESON, P. 2007. *Measuring behaviour: an introductory guide,* Cambridge, Cambridge University Press.

METZ, D. 2008. The myth of Travel Time Saving. *Transport Reviews,* 28**,** 321 - 336.

MOKHTARIAN, P. L. 2005. Travel as a desired end, not just a means. *Transportation Research Part A: Policy and Practice,* 39**,** 93-96.

NASH, J. 1975. Bus riding: community on wheels. *Journal of Contemporary Ethnography,* 4**,** 99-124.

NEW ZEALAND TRANSPORT AGENCY 2008. General Circular – Policy No. 08/09 on 'Public transport concession scheme for SuperGold Card holders'

OHMORI, N. & HARATA, N. 2008. How different are activities while commuting by train? A case in Tokyo. *Tijdschrift voor economische en sociale geografie; Royal Dutch Geographical Society* 99**,** 547-561.

PATTON, M. 2002. *Qualitative research & evaluation methods,* Thousand Oaks, SAGE.

ROBSON, C. 1993. *Real world research: a resource for social scientists and practitioner-researchers,* Oxford, Blackwell.

RUSSELL, M., PRICE, R., GERRING, Z., SIGNAL, L., CUMMING, J. & STANLEY, J. in press. What do passengers do in travel time? Structured observations on buses and trains. *Journal of Public Transportation*.

SACKETT, G. P. 1978. Measurement in observational research. *In:* SACKETT, G. P. (ed.) *Observing behavior: v 2 Data collection and analysis methods.* Baltimore: University Park Press.

SACKETT, G. P., RUPPENTHAL, G. C. & GLUCK, J. 1978. Introduction: an overview of methodological and statistical problems in observational research. *In:* SACKETT, G. P. (ed.) *Observing behavior: v 2 Data collection and analysis methods.* Baltimore: University Park Press.

SALMOND, C., CRAMPTON, P. & ATKINSON, J. 2007. NZDep2006 Index of Deprivation. Wellington: Department of Public Health, University of Otago.

STATISTICS NEW ZEALAND 2005. Understanding and working with ethnicity data: a technical paper.

THOMAS, J. 2009. *The social environment of public transport.* PhD, Victoria University of Wellington.

TIMMERMANS, H. & VAN DER WAERDEN, P. 2008. Synchronicity of activity engagement and travel in time and space: descriptors and correlates of field observations. *Transportation Research Record: Journal of the Transportation Research Board* 1-9

WARDMAN, M. 1998. The value of travel time: a review of British evidence. *Journal of Transport Economics and Policy,* 32**,** 285-316.

WARDMAN, M. 2001. Public transport values of time. Leeds: Institute of Transport Studies, University of Leeds.

WATTS, L. 2008. The art and craft of train travel. *Journal of Social and Cultural Geography* 9**,** 711-726.

WATTS, L. & URRY, J. 2008. Moving methods, travelling times. *Environment and Planning D: Society and Space,* 26**,** 860-874.