Research in Transportation Economics 69 (2018) 68–76

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Analysing the trip and user characteristics of the combined bicycle and transit mode

Sanmay Shelat[a,*], Raymond Huisman[b], Niels van Oort[a,b]

[a] Dept. of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628CN, Delft, The Netherlands
[b] Goudappel Coffeng, P.O. Box 161, 7400, AD, Deventer, The Netherlands

A R T I C L E I N F O

JEL classification:
O18
L92
R41

Keywords:
Bicycle
Transit
Bicycle-transit integration
Latent class cluster analysis
Netherlands

A B S T R A C T

Several cities around the world are facing mobility related problems such as traffic congestion and air pollution. Although limited individually, the combination of bicycle and transit offers speed and accessibility that can compete with automobiles by complementing each other's characteristics. Recognising the potential benefits with regard to accessibility, health, and sustainability, several studies have investigated policies that encourage integration of these modes. However, the actual users and trips of the combined bicycle and transit mode have not been extensively studied empirically. This study addresses this gap by (i) reviewing empirical findings on related modes, (ii) deriving user and trip characteristics of the combined bicycle and transit mode in the Netherlands, and (iii) applying latent class cluster analysis to discover prototypical users based on their socio-demographic attributes. Most trips by this combined mode are found to be for relatively long commutes where transit is in the form of trains, and bicycle and walking are access and egress modes respectively. Furthermore, seven user groups are identified and their travel behaviour is discussed. Transport authorities may use these empirical results to further streamline integration of bicycle and transit for its largest users as well as to tailor policies to attract more travellers.

1. Introduction

Cities around the world are facing several problems due to the widespread use of automobiles. Apart from problems of traffic congestion becoming commonplace, the private car is also responsible for contributing significantly to air pollution and consequent respiratory health issues. Furthermore, over-reliance on cars has lowered liveability in cities by reducing the space available for human interaction and by fragmenting the urban fabric (gtz, 2009). Having identified these car-related problems, authorities often seek to increase the modal share of active modes such as walking and cycling, and mass transit modes such as metros, bus rapid transit systems and trains (European Commission, 2011). These modes contribute to efficient, sustainable and economically vital cities (van Oort, van der Bijl, & Verhoof, 2017). However, due to their individual characteristics, they are unable to compete with automobiles: active modes have a low spatial reach due to low speed and high effort whereas transit modes, by nature, do not provide door-to-door accessibility. The car, on the other hand, is a flexible mode capable of overcoming limitations of both these modes.

As a viable competitor to cars, the combination of bicycle and transit has been found to be powerful (Brand, Hoogendoorn, van Oort, & Schalkwijk, 2017). Combined, active modes and high-level transit offer speed and accessibility in the range of personal vehicles or on-demand transit by complementing each other's characteristics (Kager, Bertolini, & Te Brömmelstroet, 2016). While transit is able to overcome the distance barrier that the bicycle faces, using bicycles as an access/egress mode significantly increases the catchment area of transit stations, thereby potentially overcoming the first and last mile hurdle that high level transit faces. This synergetic relationship between bicycles and transit has been noted by researchers, transportation institutes, and policymakers for its potential to increase sustainability, efficiency, and equity of transportation in cities around the world leading to active efforts towards the integration of these modes (Krizek & Stonebraker, 2011). These efforts include measures such as improving bicycle tracks to and from transit stations, developing and increasing bicycle parking at stations, allowing bicycles on transit vehicles and setting up public bicycle systems (gtz, 2009; Krizek & Stonebraker, 2011).

The factors affecting the use of this combined mode can be divided into four parts: (i) infrastructural facilities, (ii) policies, (iii) user characteristics, and (iv) travel characteristics. The first two parts, infrastructural facilities and policies regarding the integration of bicycle and transit, have been discussed extensively (e.g. (Krizek &

* Corresponding author.
E-mail addresses: s.shelat@tudelft.nl (S. Shelat), Hqr@goudappel.nl (R. Huisman), N.vanOort@tudelft.nl (N. van Oort).

https://doi.org/10.1016/j.retrec.2018.07.017
Received 7 December 2017; Received in revised form 10 June 2018; Accepted 16 July 2018
Available online 26 July 2018
0739-8859/

Stonebraker, 2011; Pucher & Buehler, 2009; Replogle, 1992) and (Harms, Bertolini, & Brömmelstroet, 2015; Pucher & Buehler, 2008), respectively). Few studies, however, consider the actual trips conducted or those who make the trips. Therefore, although the synergy of bicycle and transit has been recognised and several efforts towards better integration of these modes have been made, the understanding of the users and trips of this combined mode is not extensive.

To fill this scientific gap and support authorities in making decisions towards diverting more travellers to the sustainable combined bicycle and transit mode, this paper analyses the current combined mode trips and users, and identifies prototypical users of this mode. Results of this study are expected to assist policy makers with marketing strategies and infrastructure and service investment decisions that encourage modal shift.

The Netherlands is chosen for this study as extensive bicycle-related services are already present almost ubiquitously for all high level transit modes, and nearly 47% of the train passengers use a bicycle in some part of their trip (Kager et al., 2016). Moreover the Dutch Government recognises the potential of this combined mode and considers maximising its share as a policy challenge (Rijksoverheid, 2016).

After this introduction, the remainder of the paper is organized as follows. In the next section, definitions used in this study are presented. Existing empirical findings on modes related to the combined mode are reviewed in section 3. The methodology of the study is detailed in section 4, followed by the results in section 5. Finally, conclusions of the study are discussed in section 6.

2. Definitions

Since different terminologies related to the parts of a journey are found in literature, first, the terms used in this paper are defined. Four levels of a journey are considered from lowest to highest:

(i) Trip segment: Part of the journey that is performed using a single mode continuously.
(ii) Trip part: Describes the position of the travel in the journey from origin to destination: access, main, or egress. A trip part may consist of more than one trip segments.
(iii) Trip: Journey from origin to destination composed of access, main, and egress trip parts.
(iv) Trip chain: Sequence of consecutive trips starting and ending at home.

In the dataset used for the analysis (section 4.1), respondents are asked to state the main mode for each trip. The main mode is used to identify the main trip part, which also includes trip segments between those carried out by the stated main mode as enchained trip segments. Trip segments before the first main mode trip segment are for access, and those after the last main mode trip segment are for egress. Further, the mode used for the access and egress trip parts is assumed to be that which is used by the longest trip segment in the respective trip parts. This definition is reasonable, as people are unlikely to have very complex trip compositions due to the cost of transferring.

Kager et al. (2016) propose analysing the combined bicycle and transit mode as a trip chain from origin to destination since the relative attractiveness of modes is determined at this level. This also allows authorities to consider all segments of the trip to increase the combined use of bicycle and transit including complementary modes such as walking and feeder transit. Following Kager et al. (2016), this paper uses the following definition for the bicycle and transit mode:

A trip is said to make use of the bicycle and transit mode if the trip has, as its main mode, a public transport trip, in the form of a train, bus, tram, or metro, and makes use of sustainable modes: walking, cycling or feeder transit only, to access and egress it such that at least one of the access and egress trips is a bicycle trip.

This definition allows the combined mode to be seen as a truly sustainable mode without the use of private automobiles in any trip segment and consists of the following trip part combinations: (i) bicycle-transit-walk, (ii) walk-transit-bicycle, (iii) bicycle-transit-feeder transit, (iv) feeder transit-transit-bicycle, and (v) bicycle-transit-bicycle.

3. Literature review

The potential of bicycle and transit as a combined mode has been recognised in literature, with most of the work on this topic occurring in the last two decades in Western European countries like Netherlands, Denmark, Germany and the United Kingdom (Martens, 2004; Rietveld, 2000; van Mil, Leferink, Annema, & van Oort, 2018); and North American countries such as the United States and Canada (Flamm & Rivasplata, 2014; Pucher & Buehler, 2009; Tsenkova & Mahalek, 2014). This section reviews the empirical findings related to the combined bicycle and transit mode from these studies.

The variables affecting the choice of the combined bicycle and transit mode are a combination of factors affecting the use of these modes individually. Heinen, van Wee, and Maat (2010) classify determinants of bicycle use for commuting into the built environment; natural environment; socio-economic attributes; psychological factors; and aspects related to cost, effort and safety. Further, the social environment and culture regarding cycling is also an important factor in its use (Handy, van Wee, & Kroesen, 2014). Transit which is generally a part of a multimodal trip depends strongly on trip characteristics such as access/egress distance to the main transit mode (Krygsman, Dijst, & Arentze, 2004), activity purpose, duration and sequence (Susilo & Dijst, 2009); car ownership and household interactions deciding car availability; socio-demographic characteristics (Krygsman & Dijst, 2001; Nobis, 2007); and activity location and spatial characteristics (Krygsman & Dijst, 2001).

Access and egress are critical parts in multimodal trips as their importance in terms of effort or perceived time spent is much higher than their contribution in reaching the destination. The catchment area of a transit station depends (to varying degrees) on trip characteristics such as access/egress mode, transit mode or service and main trip part distance.

Intuitively, the catchment area using walking to access/egress will be smaller than cycling because of the lower speeds and greater effort in walking. Krygsman et al. (2004) find, using observed travel times, that the median access distance by cycling is 1.8 km while by walking is 550 m and the median egress distances are 2.4 km and 600 m respectively. Although they find that travellers have a fixed travel time budget for access and egress, using stated preferences of those willing to combine cycling and transit, Bachand-Marleau, Larsen, and El-Geneidy (2011) conclude that higher access/egress times on a bicycle are more acceptable to travellers than by walking. Based on a mobility survey in the Netherlands, Keijer and Rietveld (2000) calculate the share of different modes for different distance classes and find that walking dominates access trips up to 1.5 km and egress trips up to 2.5 km. The switch to bicycles at 1.5 km at the home-end, and the fact that at the activity end feeder transit is used more commonly for distances too great to walk indicate the effect of bicycle availability on mode choice. Based on the fact that most access and egress trips are up to 2.5 km, walking and cycling are the most common access and egress modes to transit stops.

The type of main mode also affects the acceptable access and egress distance (Daniels & Mulley, 2013). It is evident that services likely to have a higher frequency have a larger catchment area indicating that people are willing to travel further to reach a more frequent mode (Alshalalfah & Shalaby, 2007). For example, Brand et al. (2017) demonstrate that access and egress distances to stops of high quality transit lines are about twice as large as those of regular transit. Reliability and level of service are also factors in the acceptable access distance, for example O'Sullivan and Morrall (1996) find that LRT stations in Calgary, Canada have almost double the catchment radius than suggested by standard bus guidelines.

Longer trip distances encourage the use of public transportation (Keijer & Rietveld, 2000) and consequently the accessibility radius also increases. However, the access and egress times increase at a lower rate than the main segment therefore the proportion of access and egress time out of total trip time (or interconnectivity ratio) decreases with longer trip times (Krygsman et al., 2004).

The series of activities travellers would like to perform also has an impact on their choice of mode, for example, if a traveller has to go to work in the morning and also drop of their children to school they are more likely to use a car because of its higher capacity and also because the number of transfers would be very high with public transportation. Hensher and Reyes (2000) postulate that higher complexity of trips lowers the likelihood of using public transportation. However, they find that this is true only for work trips and not for non-work trips and that car availability and households with children form barriers to the use of public transportation. The combined mode is used mostly for daily activities (such as work and education) rather than incidental activities such as shopping (Martens, 2004). However, when the *OV-fiets* (Dutch public bicycle renting scheme) is used as egress mode it is typically for more infrequent activities where people may not have a second bicycle (Martens, 2007).

Among the socio-demographic variables, those related to transportation characteristics have been found to be the most important in the decision to use multimodal travel such as the combined bicycle and transit mode (Krygsman et al., 2004; Kuhnimhof, Chlond, & Huang, 2010). These include attributes such as number of cars in the household, possession of license, public transport usage, possession of a public transport card, etc. Those who do not own a car or do not have access to one are called captive travellers as they are forced to use other modes. Krygsman et al. (2004) find that 80% of multimodal transport users, who are a very relevant group for the bicycle + transit mode, do not own a personal car and Kuhnimhof et al. (2010) have similar findings regarding car owners' bicycle use within the radius of non-motorized travel. In the multimodal bicycle travellers cluster of Molin, Mokhtarian, and Kroesen (2016), p. 38% never have access to a car while 26% have some difficulty in arranging for a car and similar traveller clusters in Nobis (2007) and Anable (2005) are also characterised by low car availability. The acceptable access distance for those who have a license has been found to be higher than those who don't (Alshalalfah & Shalaby, 2007). A similar effect is also observed for the possession of transit pass with those who own a pass travelling a lower distance to access public transportation. These effects can be attributed to self-selection of residences wherein those who expect to use transit – because they do not have a license and have a pass – live nearby transit stops.

Cluster analysis, an exploratory statistical technique, used to form groups of observations different from one another but similar within themselves has been used in several studies on travel attitudes and multimodality. A typical method is to first conduct a factor analysis on variables to reduce their number by combining correlated variables into the same factors so that they do not have an inordinate effect on the clustering. This is especially important in attitudinal studies where the number of attitudinal variables is quite large. Next, the regression scores of each observation are used as input to a clustering algorithm such as k-means. Anable (2005) use this method to understand the mode-switching potential of travellers in the UK and create traveller clusters with unique psychographic profiles. Olafsson, Nielsen, and Carstensen (2016) identify multimodal travel patterns in the Danish population using this method and examine the role of cycling in these patterns. Applying this analysis on responses of a questionnaire designed specifically to understand how travellers wish to integrate bicycle and transit, Bachand-Marleau et al. (2011) identify 5 types of travellers which they further classify into current, potential, and improbable integrators of the two modes.

Amongst latent construct based methods, after reducing the attitudinal variables with a factor analysis, Shiftan, Outwater, and Zhou (2008) employ structural equation modelling to obtain relationships between the reduced traveller attitudes and socio-demographic data, and between these and latent attitudinal factors. Finally, attitudinal factor scores are used to group people with similar attitudes together to identify potential transit markets. Molin et al. (2016) use latent class cluster analysis to identify modality patterns and attitudes of travellers. Several studies operationalize these latent constructs in a choice modelling context through hybrid choice models. The aim of these models is to reduce the unobserved variations in behaviour by taking into account latent preferences caused by attitudes, beliefs, and habits. Estimating such models for the binary choice of bicycle or not as the access mode choice to train stations, La Paix Puello and Geurs (2015) find that both observable variables as well as attitudes are important. Handy et al. (2014) provide a summary of advanced statistical techniques that can be applied to transportation mode use research.

While the above studies provide many insights, there is potential to add to this knowledge through empirical analysis of the users and trips of a well-defined combined bicycle and transit mode and identification of its prototypical users. The next section will present our approach.

4. Methodology

In order to analyse the trip and user characteristics of the combined mode, data from a national survey in the Netherlands is used. In this section, first, the survey sample and dataset structure is described, followed by an overview of the trips in the dataset and a discussion on the representativeness of the extracted combined mode trips. The second sub-section presents the clustering methodology used and applies it on the extracted combined mode users to identify prototypical users.

4.1. Dataset description and extraction of combined mode trips

To gain insights into the current characteristics of bicycle and transit users, the Dutch national one-day trip diary survey, *Onderzoek Verplaatsingen in Nederland* (OViN), from 2010 through 2015 is used (Centraal Bureau voor de Statistiek, 2011–2016). Sampling for OViN is done so that the data is representative of population at the level of the 12 provinces of the Netherlands. Moreover, the survey has a large sample size with over 250,000 persons interviewed in the six years considered for this study.

The sample unit of OViN is an individual (as opposed to a household) and each response contains information about the individual's personal and household characteristics as well as details of all the trips they make on a particular day. Hence, variables in the OViN dataset can be broadly divided into socio-demographics and travel characteristics. Table 1 partitions the main variables used or considered in the analysis into these categories.

Each record in the OViN dataset represents one trip segment, such that, a group of consecutive records make up a trip, then a trip chain, and, finally, all the travel carried out by the respondent. In each record, the socio-demographic characteristics of the respondent is repeated. To analyse the different trips and users, first, the dataset is partitioned into a socio-demographic dataset and travel-related dataset. The travel-related dataset is manipulated to form three datasets in which each record is, respectively, a trip segment, a trip, and a trip chain.

Next, trips made by the combined bicycle and transit mode are extracted. Since, a trip by the combined bicycle and transit mode is, by definition, a multimodal trip, first, all the multimodal trips are extracted. Any trip with more than one trip or with transit as the main mode is defined as a multimodal trip. For trips with transit as the main mode, if the access or egress trip part is missing, it is assumed that the respondent considers this as a ‘very short walk’. From these multimodal trips, those with transit as the main mode, and sustainable access and egress modes (walk, bicycle, or transit) with at least one being bicycle,

Table 1
Variables in the OViN dataset used in the analysis.

Socio-demographic	Travel related
Household	*Overall*
• Number of persons in the household • Household composition • Degree of urbanization of residence • Disposable household income	• Reporting date and day of week • Total travel duration in the day • Total distance travelled in the day
Individual	*Trip*
• Gender • Age • Mode of social participation • Highest education	• Trip has the same origin and destination • Number of trip segments • Trip destination • Trip motive • Departure and arrival postcodes • Trip distance • Main mode used in trip • Departure and arrival times • Trip travel time • Activity duration
Transportation	*Trip segment*
• Number of cars in the household • Driving license • Main user of a car in household • Transit use frequency • Student Transit pass availability	• Trip segment distance • Trip segment mode • Trip segment departure and arrival times • Trip segment travel time

Table 2
Overview of records in each dataset.

Dataset	Conditional Filter	Number of records	%age relative to overall
OViN (Trip segments)		806,011	–
Trips		684,245	–
Trip chains		298,898	–
Socio-demographic (Users)		252,110	–
Multimodal trips	Trip segments > 1	35,466	5.2% (of all trips)
Multimodal users		19,620	7.8% (of all users)
Combined mode trips	See mode definition	5943	0.9% (of all trips)
Combined mode users		3376	1.3% (of all users)

are selected. These make up the trips made by the combined mode. Any user that makes at least one multimodal trip is called a multimodal user, and, similarly, any user that makes at least one trip with the combined mode is called a combined mode user. Table 2 provides a summary of records in the different datasets.

Despite the large survey sample and representative sampling in the OViN dataset, selective non-responses may lead to certain sections of the population being over- or under-represented. To correct this, and enable scaling the sample to the population, weighting factors are calculated based on certain background variables whose true or better estimated distributions are known. However, the representativeness of these weighting factors deteriorate as more conditions are applied to filter the dataset.

This is relevant here, because the definition of the combined mode requires the application of multiple conditional filters on the data to extract trips made using it. As a result of these conditional filters, as shown in Table 2, the number of combined mode users in the sample is quite small. Therefore, in the analysis, the combined mode sample is not scaled to the population, as the representativeness after weighting is not expected to be much better than the sample itself. The addition of more data over the next years should allow the sample to remain representative even after the application of multiple filters.

4.2. Clustering of combined mode users

Cluster analysis is a type of unsupervised learning that uses unlabelled data to classify multivariate data into natural groups such that the observations within a cluster are highly similar to each other but are very dissimilar to observations in other clusters (Han, Kamber, & Pei, 2006; Izenman, 2008, pp. 237–280). Since the objective in this chapter is to identify prototypical users or, to put it differently, classify combined mode users into natural groups, cluster analysis is very useful. Clustering is used in a wide range of studies but here the objective is similar to that in marketing studies where this technique is used to segment markets into small homogeneous groups so that targeted campaigns can be carried out efficiently (Izenman, 2008, pp. 237–280).

Latent class cluster analysis (LCCA) is chosen as the clustering algorithm for this study. Unlike standard clustering techniques which are based on partitioning or hierarchical methods, LCCA is a model-based method. Since LCCA is not based on proximity measures, it can take as input both categorical and numerical variables. This is useful because the dataset used here contains a mixture of data types; for example, categorical variables such as occupation, and ordinal variables such as degree of urbanization of residence. Furthermore, being a statistical model, LCCA provides formal statistical criteria to decide the number of clusters which is typically difficult for other clustering methods (Vermunt & Magidson, 2002).

In LCCA, the basic idea is that a categorical latent variable can account for the covariation between different indicator attributes of observations. The categories of this latent variable are the clusters of which observations (in this case, combined mode users) are probabilistic members. The aim of the analysis is to find the least number of clusters (or categories) that render the covariation between the indicators insignificant. Interested readers are guided to (Vermunt & Magidson, 2002) for more detailed formulation of LCCA and to (Molin et al., 2016) for a shorter summary. This study uses LatentGOLD 5.1 (by Statistical Innovations), a commercial software, for applying LCCA.

Since the objective of applying LCCA here is to find prototypical users of the combined bicycle and transit mode, the indicator variables all consist of socio-demographic characteristics (Table 1) while there are no covariates. The variable, number of persons in the household, is not used as this has a very high correlation with the household composition. Further, the variables, number of cars in household, main car user, and driving license have been combined into a single variable – car availability. Finally, student transit pass availability variable is removed because of its high correlation with age and mode of social participation. All indicator variables and their categories can be seen in Table 4.

Table 3
Model fit of 1–10 cluster LCCA models.

No. of clusters	np	LL	df	L^2	p-value	BIC
1	50	−41823.4	3326	31065.48	0.00	84053.08
2	101	−38192.2	3275	23802.96	0.00	77204.9
3	152	−36595	3224	20608.54	0.00	74424.83
4	203	−35650.7	3173	18720.03	0.00	72950.67
5	254	−35212.3	3122	17843.32	0.00	72488.31
6	305	−34897.2	3071	17212.96	0.00	72272.3
7	**356**	**−34667.5**	**3020**	**16753.67**	**0.00**	**72227.35**
8	407	−34528	2969	16474.72	0.00	72362.75
9	458	−34421.5	2918	16261.65	0.00	72564.02
10	509	−34345.3	2867	161015	0.00	72825.87

np – number of parameters.
LL – log-likelihood.
df – degrees of freedom.
L^2 – Likelihood-ratio chi-square statistic.

Table 4
Profile of the 7 cluster model.

Cluster number	1	2	3	4	5	6	7
Cluster size	26.50%	22.84%	15.10%	14.06%	9.64%	9.47%	2.39%
Label	Middle-aged full-time professionals	University students living with parents	School children	Young, low income professionals	Middle-aged part-time professionals	University students living alone	Pensioners
Household indicators							
Household composition							
Single	4%	1%	1%	**64%**	2%	**58%**	**39%**
Couple	**37%**	1%	0%	**24%**	**31%**	11%	**59%**
Couple + kid(s)	**55%**	**84%**	**86%**	1%	**64%**	2%	0%
Couple + kid(s) + other(s)	0%	1%	0%	0%	1%	0%	0%
Couple + other(s)	0%	0%	0%	0%	0%	1%	0%
Single parent + kid(s)	2%	10%	12%	8%	3%	3%	1%
Single parent + kid(s) + other(s)	0%	0%	0%	0%	0%	1%	0%
Other composition	1%	2%	0%	2%	0%	25%	1%
Degree of urbanization							
Extremely urbanised	**28%**	7%	6%	**50%**	**21%**	**61%**	**23%**
Strongly urbanised	**36%**	**22%**	19%	**30%**	**33%**	**29%**	**25%**
Moderately urbanised	**22%**	**26%**	**21%**	13%	**24%**	6%	**34%**
Hardly urbanised	11%	**30%**	**25%**	4%	14%	3%	9%
Not urbanised	4%	15%	**29%**	2%	8%	1%	10%
Household disposable income							
Lowest - 1st 10%ile	1%	5%	7%	9%	5%	**71%**	3%
2nd 10%ile	2%	4%	**10%**	**10%**	5%	10%	**10%**
3rd 10%ile	3%	7%	**9%**	**10%**	5%	5%	7%
4th 10%ile	7%	9%	**12%**	**13%**	7%	1%	**12%**
5th 10%ile	6%	10%	**9%**	**13%**	10%	2%	**10%**
6th 10%ile	11%	**13%**	**10%**	**10%**	11%	3%	**10%**
7th 10%ile	**12%**	**14%**	**11%**	**12%**	**15%**	1%	**11%**
8th 10%ile	**18%**	**17%**	**10%**	9%	**12%**	1%	**12%**
9th 10%ile	**21%**	**14%**	**12%**	6%	**17%**	3%	8%
Highest - 10th 10%ile	**20%**	7%	**9%**	6%	**13%**	2%	**16%**
Unknown	0%	0%	0%	1%	0%	1%	0%
Individual							
Gender							
Male	**70%**	52%	47%	42%	15%	39%	45%
Female	30%	48%	53%	58%	**85%**	61%	55%
Age group							
1 to 17	0%	1%	**100%**	0%	0%	0%	0%
18 to 24	2%	**98%**	0%	4%	5%	**84%**	0%
25 to 34	24%	1%	0%	**49%**	17%	15%	0%
35 to 64	**74%**	0%	0%	**46%**	**78%**	1%	12%
> =65	0%	0%	0%	0%	0%	0%	**88%**
Highest education							
None	0%	1%	2%	0%	0%	0%	0%
Primary	0%	2%	**25%**	1%	0%	0%	1%
Secondary	3%	19%	**26%**	7%	9%	3%	16%
High school	19%	**70%**	**18%**	24%	**29%**	**55%**	19%
University	**76%**	7%	0%	**67%**	**62%**	**41%**	**63%**
Other	1%	1%	1%	1%	0%	1%	0%
Unknown	0%	0%	0%	0%	0%	0%	1%
N/A (< 15 years)	0%	0%	29%	0%	0%	0%	0%
Social participation							
Work (12–30h)	0%	6%	3%	11%	**67%**	14%	4%
Work (> =30h)	**99%**	7%	2%	**81%**	18%	9%	0%
Homemaker	0%	0%	0%	1%	5%	0%	0%
Student	0%	**86%**	**94%**	0%	1%	**75%**	0%
Unemployed	0%	1%	0%	2%	1%	2%	0%
Disabled	0%	0%	0%	3%	1%	0%	0%
Pensioner	0%	0%	0%	0%	0%	0%	**95%**
Other	0%	0%	0%	3%	6%	1%	0%
N/A (< 6 years)	0%	0%	2%	0%	0%	0%	0%
Transport							
Car availability							
Other	0%	0%	0%	0%	0%	0%	0%
Always available	43%	10%	0%	21%	28%	6%	42%
Limited availability	46%	49%	2%	0%	52%	7%	23%
No car in household	5%	0%	0%	49%	4%	59%	11%
Younger than 18	0%	0%	98%	0%	0%	0%	0%
No license to drive	6%	41%	0%	30%	16%	28%	24%
Transit use frequency							
Daily	**60%**	**70%**	**66%**	**53%**	14%	**41%**	3%
Once a week	**32%**	27%	**19%**	**33%**	**65%**	**51%**	**52%**
Once a month	6%	2%	8%	13%	13%	7%	**34%**
< Once a month	3%	0%	5%	1%	7%	0%	11%
Almost never	0%	0%	2%	0%	1%	1%	0%

LCCA offers different statistical criteria to identify the number of clusters required. Two commonly used criteria are the global and relative model fits. The global model fit tests, using the chi-square statistic, whether the observed and expected values of different response patterns within each class is the same (that is, indicator variables do not have any association). A response pattern is a particular set of responses from all the indicators; for example, the set [male, single household, 2nd percentile disposable income] is one possible response pattern for the indicators [gender, household composition, disposable income]. Since many indicator variables – each consisting of multiple categories – are used here, the number of possible response patterns is quite large. Therefore, the chi-square statistic is unlikely to follow the chi-square distribution in such sparse data. In the relative model fit criterion, the idea is to balance the model fit (log-likelihood) and model parsimony (number of parameters) as more clusters would always lead to a better model fit. The Bayesian information criterion (BIC) is used for this – the lowest BIC value indicates the best fitting model. This method is used here as it can also be applied to situations with low number of responses per response pattern.

In order to decide the number of clusters, the models for 1 to 10 clusters are estimated (Table 3) and the bivariate residuals (not shown) and BIC values examined to determine the number of clusters that should be used. As can be seen, the chi-square statistic, L^2, rejects even the 10 cluster model (the statistic should not reject the null hypothesis to show that the observed and expected values are the same). All the bivariate residuals are first insignificant (free of any significant associations between indicator variables) at the 7 cluster model. The lowest BIC value is also at the 7 cluster model. Moreover, upon examination of the clusters, they were found to easily interpretable. Further, each cluster except the smallest has more than 100 members indicating the reliability of the results.

5. Results

This section presents our main findings with regard to the descriptive statistics on the trip and user characteristics and clustering of the combined mode users. The first section is split into three parts, the first part shortly deals with findings for main travel modes while the second deals with access/egress modes. Findings for the combined bicycle and transit mode are discussed last. The second section of this chapter describes the user characteristics of prototypical combined mode users in further detail.

5.1. Trip and user characteristics

5.1.1. Main modes

Fig. 1 shows the (main) mode share by distance classes. Walking is the preferred mode for short distances up to 1 km after which cycling is the dominant mode for distances up to 5 km. At the opposite spectrum of distances, trains are used substantially only for distances larger than 10 km. Other transit, however, which includes bus, tram, and metro, is used for a large range of distances (3–40 km). This can be expected as these modes range across multiple transit network levels.

Besides distances, travel time is an important indicator of connectivity. Cycling and walking have comparable travel time decays; however, due to their higher speed, cyclists cover a larger distance within the same acceptable time. This indicates that, compared to walking, when cycling is used as access and egress mode for transit, a larger catchment area exists.

In terms of socio-economic variables, car and train are used more by high income users while bus, tram, and metro have more low income users (Fig. 2). Transit modes have a large proportion of users from the lowest income group. Dutch students living alone (with no income of their own) may have contributed to this large proportion as they receive free transit passes for the duration of their study. Unlike the above modes, bicycle users can be found across the income range.

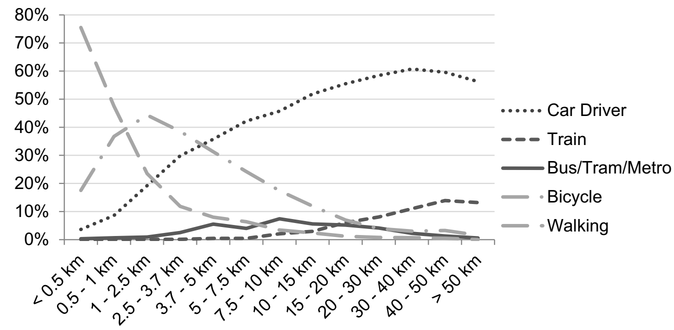

Fig. 1. Mode shares of car, train, bus/tram/metro, and walking per distance classes.

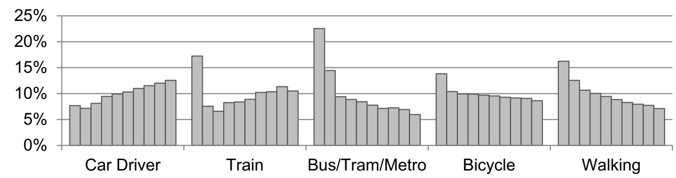

Fig. 2. Household disposable income (standardised 10th percentiles; lowest - left) distribution per mode.

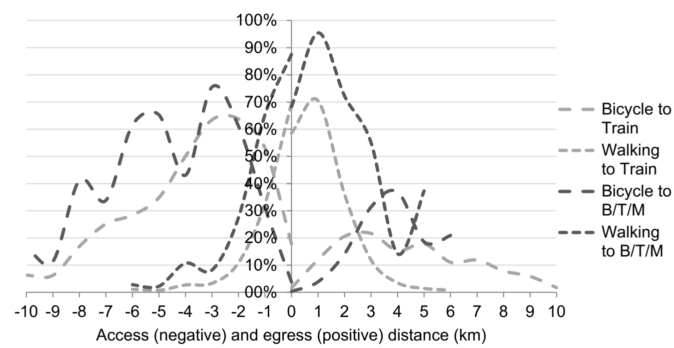

Fig. 3. Mode shares per distance class of cycling and walking in access and egress trips for train and bus/tram/metro (B/T/M) stations.

5.1.2. Access/egress modes

Since access and egress trips are symmetric by nature – what is access in a home-based trip becomes egress in an activity-based one –, here all access and egress trips occur at the home and activity ends, respectively. Transit users are willing to travel larger distances towards train stations than bus, tram, or metro station with a mean home-end access distance of 3.8 km versus 1.5 km, respectively. This order also holds for activity-end egress distance but the distances are smaller here with egress distances 2.7 km and 0,7 km, respectively. Smaller activity-end distances are expected as popular activity destinations are likely to have a transit station nearby. Moreover, since travellers have fewer egress mode options at the activity end, they are likely to get off the main mode as close to their destination as possible.

Fig. 3 shows the distance decay curves for access and egress modes to transit stations. The squiggly nature of the curve for bus/tram/metro stations may be attributed to rounding off in the self-reported distances and the smaller number of responses. It can be seen that bicycle substitutes walking as the most popular access and egress mode at a lower distance for trains than for lower level transit networks. Thus, it can be seen that travellers are either reluctant to combine cycling with bus/tram/metro or do not have the facilities to do so.

Access distance increases with the main mode distance for both train and bus, tram, and metro trips; although, the rate of increase is higher for bus, tram and metro than train. A positive relation is also

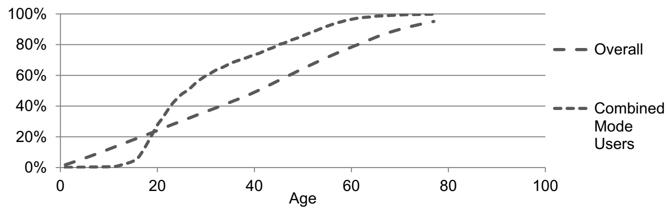

Fig. 4. Cumulative distribution of combined mode users' age.

found between the frequency of transit use and the use of the bicycle as an access mode. This is mainly due to the fact that the bicycle and transit mode is often used for work, business and education purposes which are likely to be activities done more than once a week.

5.1.3. Combined bicycle and transit mode

The average distance of a trip by the combined bicycle and transit mode is 41 km which is longer than the average multimodal trip, indicating that the combined mode may be more suitable for long trips. The average distance is larger when the combined model includes transit as a feeder function; although, this is not the most common trip combination. The most common combination is the bicycle-transit-walk (56%). This combination is followed by bicycle-transit-feeder transit (23%), bicycle-transit-bicycle (15%), and walk-transit-bicycle and feeder transit-transit-bicycle both at 3%. Furthermore, 82.8% of the transit modes within combined mode trips is 'train', whereas the remaining 17.2% consist of bus, tram or metro as the main mode. The majority of trips are used to go to work (51%) or education (31%). Moreover, most combined mode trips either start from or end at home. The trip objectives imply that most trips are made multiple times per week and mostly on weekdays during morning rush hours.

Regarding the socio-demographics of the combined mode users, virtually no differences exist in the gender distribution from the overall distribution and both men and women are equally represented. However, users are more likely to belong to a higher household income group and be more educated (nearly 85% have at least high school education) than the overall sample. A possible reason for this may be that high educated, high income users have specialised jobs that require commuting longer distances for which the combined mode is suitable. Finally, as shown in Fig. 4, ages between, approximately, 17–27 are over-represented amongst the combined mode users while those younger than 12 and older than 60 are under-represented.

5.2. Clustering bicycle and transit users

The OViN dataset is used to identify prototypical bicycle and transit users by defining a 7-cluster model using LCCA. The profile of the 7 cluster model is shown in Table 4. Subjective estimates of where the bulk of the values in each indicator per class lies is shown in bold. Based on these properties of the clusters, each cluster is given a label defining the prototypical user represented by that cluster. It should be noted that this label is a subjective, average group definition and does not imply that all bicycle and transit users belonging to a cluster have the properties of the label. Next, the properties of each cluster is discussed.

The smallest cluster, **pensioners** represent just under 2.4% of the total group of bicycle and transit users. Although the number of members in this group is quite low, the group has been added because their properties are too different from the other groups. Most of the members are highly educated, and approximately 40% of these travellers always have access to a car. Further reinforcing the idea that this group is not very important for the bicycle and transit mode is the fact that they are not regular transit users and most of them use transit either once every week or month. More than half of the users live in extremely or strongly urbanised regions which is expected as these occasional travellers are likely to use the combined mode only if it is easily possible as in urban areas where transit level of service and bicycle path density are both higher.

In terms of increasing group size, the pensioners are followed by **university students living alone** (9.5%). A large portion, around 20%, uses this mode to visit someone. Like the part-timers of the next group (cluster 5), the share of those using transit daily is less than half. This means that a large share of travellers do not use the bicycle and transit mode daily. Moreover, their trips are spread out over the day unlike most other clusters. Like the young, low income professionals (cluster 4) nearly 90% of this group's members never have access to a car and are therefore captive users of the bicycle and transit mode. Nearly all of this group's members live in highly urbanised areas – most of them either in a university city or in a nearby city from where they can reach the university easily.

The third group, **middle-aged, part-time professionals** represent more or less the same size (9.6%) as cluster 6. This group is very similar to the first cluster, high-income full time professionals, in terms of income, household composition, age, and residence location. However, unlike cluster 1, 85% of the members of this group are female, and three quarters of the members do not always have access to a car. Furthermore, as the title of this cluster suggests a majority of the members are part-time professionals. Perhaps owing to the nature of part-time work, in the distribution of departure times, the evening peak is more spread out as compared to full-time workers of cluster 1.

The remaining four groups are all larger than 10% of the total group of bicycle and transit users. 14% of the users are **young professionals with a low income**. This is one of the groups, alongside cluster 1, comprising of full-time workers, but the two groups have several differences. The members of this group are younger, earn lesser and have an almost equal distribution of gender. Moreover almost nobody in this cluster has children although nearly a quarter lives with a partner. Although, the destination of their trips are also to urbanised regions, unlike cluster 1, 50% of this group also lives in extremely urbanised regions. About 80% of these travellers never have access to a car, making them captive users of this combined mode. The unavailability of cars may be the reason behind this group using the combined mode 35% of the time for purposes other than work.

The cluster **school children** (15.1%) consists of young persons (with an average age of 15), equally divided by gender. They, naturally, have no access to cars and live with their parents. The households in this group are in rural areas and have their income is on the lower side. A possible reason for the existence of this group may be that rural areas do not have schools that are close enough to be walked or cycled to and, therefore, require some form of transit modes. The average main mode distance is 25 km, lower than the average by a little more than 10 km. Furthermore, nearly 40% of the trips made by this cluster use bus, tram, or metro as the main mode which is more than double the average of all combined mode trips. 70% of the trips use walking as the egress mode while almost 20% of users cycle to their final destination.

By far the largest groups are **university students living with parents** (22.8%) and middle aged, male professionals (26.5%). To start with the penultimate one: university students living with their parents are equally represented by men and women and come from all strata of household income groups. Unlike students living alone, households of the members of this cluster are spread over semi-rural to strongly urbanised regions. Further, as one would expect, most of them have no or limited access to the household car and are, therefore, daily transit users. About a quarter of these travellers use bus, tram, or metro as their main mode, higher than the combined mode average, and significantly less (10% less) people use the bicycle at the egress end, instead replacing it with feeder transit and walking. The mean distances for the three trip parts do not differ much from the average of all clusters. Similar to university students living alone, departure times of trips are quite spread out although this group has substantially fewer trips towards the end of the day.

The largest group, finally, are **middle-aged, full time**

professionals. More than a quarter of the bicycle and transit mode users belong to this group. Most of the members (74%) of this group are middle-aged (35–64 years) working men (70%), highly educated and from high income households. Nearly all of them live with a partner while almost half have kids. They are generally (90%) one of the core members of the household, and nearly half the times, the household's main car user. This means that even though a car is available to these travellers they choose to use the bicycle and transit mode. Most of the members of this group live in urban and suburban regions of the country, and travel to extremely or strongly urbanised areas with the purpose work.

6. Conclusions

The bicycle and transit mode is attracting attention from research and policy-makers alike from around the world as a sustainable mode that is able to compete with the car. Although there is a growing interest in this mode and its benefits, not all its dimensions have been studied very well. Several research articles, guides and manuals exist on the policies and methods suitable for the integration of the bicycle and transit but literature addressing its usage characteristics is still quite rare. From the different factors affecting the use of the bicycle and transit mode, the trip and user characteristics are quite important. Knowing which types of travel and for whom the combined bicycle and transit mode is suitable is likely to be very helpful in designing policies and services that encourage the use of this mode. With this motivation, this study analysed the user and trip attributes of the bicycle and transit mode in the Netherlands.

This study presents research on combined bicycle and transit mode, consisting of (i) reviewing empirical findings on related modes, (ii) deriving user and trip characteristics of the bicycle and transit mode in the Netherlands, and (iii) applying latent class cluster analysis to discover prototypical users based on their socio-demographic attributes. Most trips by this mode are found to be for relatively long commutes to urbanised regions with transit in the form of trains, and bicycle and walking as access and egress modes respectively. The users of the combined mode are typically highly educated, have higher incomes, and are over-represented in the 17–27 age group. Amongst these users, seven prototypical groups were identified and their travel behaviour was discussed.

Our results indicate potential in increasing the market share of the combined mode, especially by improving the (activity end) egress trip part. Egress trips have a smaller average distance that could be extended by increasing the availability of bicycles. Although commuters may have a second bicycle at their activity end, increased bicycle availability would potentially attract many irregular trip purposes such as recreation or visiting somebody. New bike rental schemes, both docked and dock-less, will play a major role here. The properties of the clusters of combined mode users could help identify which travellers can be targeted as potential customers of these rental schemes. For instance, of all the combined mode trips made by university students living alone, 20% are used to visit somebody; and young professionals who do not have a car make about 35% of their combined mode trips for purposes other than work.

Although most combined bicycle and transit mode trips are made with train as the main mode, there seem to be many opportunities to increase trips with bus, tram, and metro systems. Distance decay curves of access and egress trips to stations of these modes showed that travellers switched from walking to cycling at larger distances than train stations; indicating reluctance to combine the bicycle with these modes probably due to lack of facilities. Since access and egress trip distance is positively correlated to both larger main mode distance and frequency of the main mode, combination opportunities are especially greater with high level bus, tram, and metro services. Thus, an integrated design of bicycle infrastructure (both paths and station parking facilities) and these transit services (including their stopping distance and level of service) is needed. Better combination of the bicycle and bus, tram, and metro modes can be especially marketed to university students living with their parents and school children who use this combination for 25% and 40% of their combined mode trips, respectively.

Thus, by studying the trip and user characteristics of the combined bicycle and transit mode, and identifying prototypical users, this paper helps authorities in taking actions towards increasing the share of this mode, and shift passenger transportation to a sustainable future.

Declaration of interest

None.

References

Alshalalfah, B. W., & Shalaby, A. S. (2007). Case Study: Relationship of walk access distance to transit with service, travel, and personal characteristics. *Journal of Urban Planning and Development, 133*(2), 114–118.

Anable, J. (2005). 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. *Transport Policy, 12*(1), 65–78.

Bachand-Marleau, J., Larsen, J., & El-Geneidy, A. (2011). Much-anticipated marriage of cycling and transit. *Transportation Research Record: Journal of the Transportation Research Board, 2247*, 109–117.

Brand, J., Hoogendoorn, S., van Oort, N., & Schalkwijk, B. (2017). Modelling multimodal transit networks integration of bus networks with walking and cycling. *2017 5th IEEE international conference on models and technologies for intelligent transportation systems (MT-ITS)* (pp. 750–755). .

Centraal Bureau voor de Statistiek (2011-2016). Onderzoek verplaatsing in Nederland - OViN. *DANS*.

Daniels, R., & Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use, 6*(2), 5–20.

European Commission (2011). *White Paper on transport — roadmap to a single European transport area*.

Flamm, B., & Rivasplata, C. (2014). Public transit catchment areas. *Transportation Research Record: Journal of the Transportation Research Board, 2419*, 101–108.

gtz (2009). *Cycling-inclusive policy development: A handbook*.

Handy, S., van Wee, B., & Kroesen, M. (2014). Promoting cycling for Transport: Research needs and challenges. *Transport Reviews, 34*(1), 4–24.

Han, J., Kamber, M., & Pei, J. (2006). *Data Mining: Concepts and techniques* (3 ed.). Elsevier, Morgan Kaufmann Publishers.

Harms, L., Bertolini, L., & Brömmelstroet, M. T. (2015). Performance of municipal cycling policies in medium-sized cities in the Netherlands since 2000. *Transport Reviews, 36*(1), 134–162.

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by Bicycle: An overview of the literature. *Transport Reviews, 30*(1), 59–96.

Hensher, D. A., & Reyes, A. J. (2000). Trip chaining as a barrier to the propensity to use public transport. *Transportation, 27*(4), 341–361.

Izenman, A. J. (2008). *Linear discriminant analysis, modern multivariate statistical Techniques: Regression, classification, and manifold learning.* New York, NY: Springer New York237–280.

Kager, R., Bertolini, L., & Te Brömmelstroet, M. (2016). Characterisation of and reflections on the synergy of bicycles and public transport. *Transportation Research Part A: Policy and Practice, 85*, 208–219.

Keijer, M. J. N., & Rietveld, P. (2000). How do people get to the railway station? The Dutch experience. *Transportation Planning and Technology, 23*(3), 215–235.

Krizek, K. J., & Stonebraker, E. W. (2011). Assessing options to enhance bicycle and transit integration. *Transportation Research Record: Journal of the Transportation Research Board, 2217* 1622–1167.

Krygsman, S., & Dijst, M. (2001). Multimodal trips in The Netherlands: Microlevel individual attributes and residential context. *Transportation Research Record: Journal of the Transportation Research Board, 1753*, 11–19.

Krygsman, S., Dijst, M., & Arentze, T. (2004). Multimodal public transport: An analysis of travel time elements and the interconnectivity ratio. *Transport Policy, 11*(3), 265–275.

Kuhnimhof, T., Chlond, B., & Huang, P.-C. (2010). Multimodal travel choices of bicyclists. *Transportation Research Record: Journal of the Transportation Research Board, 2190*, 19–27.

La Paix Puello, d. L. C., & Geurs, d.i. K. T. (2015). Modelling observed and unobserved factors in cycling to railway stations: Application to transit-oriented-developments in The Netherlands. *European Journal of Transport and Infrastructure Research, 15*(1), 27–50.

Martens, K. (2004). The bicycle as a feedering mode: Experiences from three european countries. *Transportation research Part D: Transport and Environment, 9*(4), 281–294.

Martens, K. (2007). Promoting bike-and-ride: The Dutch experience. *Transportation Research Part a, 41*, 326–338.

van Mil, J., Leferink, T. S., Annema, J. A., & van Oort, N. (2018). *Insights into factors affecting the combined bicycle-transit mode, CASPT 2018, Brisbane, Australia*.

Molin, E., Mokhtarian, P., & Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. *Transportation research Part A: Policy and Practice, 83*, 14–29.

Nobis, C. (2007). Multimodality: Facets and causes of sustainable mobility behavior. *Transportation Research Record: Journal of the Transportation Research Board, 2010*,

35–44.
O'Sullivan, S., & Morrall, J. (1996). Walking distances to and from light-rail transit stations. *Transportation Research Record: Journal of the Transportation Research Board, 1538*, 19–26.
Olafsson, A. S., Nielsen, T. S., & Carstensen, T. A. (2016). Cycling in multimodal transport behaviours: Exploring modality styles in the Danish population. *Journal of Transport Geography, 52*, 123–130.
van Oort, N., van der Bijl, R., & Verhoof, F. (2017). The wider benefits of high quality public transport for cities. *European transport conference, barcelona, Spain*.
Pucher, J., & Buehler, R. (2008). Making cycling irresistible: Lessons from The Netherlands, Denmark and Germany. *Transport Reviews, 28*(4), 495–528.
Pucher, J., & Buehler, R. (2009). Integrating bicycling and public transport in North America. *Journal of Public Transportation, 12*(3), 79–104.
Replogle, M. (1992). Bicycle access to public transportation: Learning from abroad. *ITEA Journal,* 15–21.
Rietveld, P. (2000). The accessibility of railway stations: The role of the bicycle in The Netherlands. *Transportation Research Part D: Transport and Environment, 5*(1), 71–75.
Rijksoverheid, 2016. https://www.rijksoverheid.nl/onderwerpen/fiets/inhoud/fietsbeleid.
Shiftan, Y., Outwater, M. L., & Zhou, Y. (2008). Transit market research using structural equation modeling and attitudinal market segmentation. *Transport Policy, 15*(3), 186–195.
Susilo, Y., & Dijst, M. (2009). How far is too Far? Transportation research record. *Journal of the Transportation Research Board, 2134*, 89–98.
Tsenkova, S., & Mahalek, D. (2014). The impact of planning policies on bicycle-transit integration in Calgary. *Urban Planning and Transport Research, 2*(1), 126–146.
Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars, & A. McCutcheon (Eds.). *Applied latent class analysis* (pp. 89–106). Cambridge: Cambridge University Press.