Multimodal Transportation 2 (2023) 100095

Contents lists available at ScienceDirect

Multimodal Transportation

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

Full Length Article

Persona Design Methodology for Work-Commute Travel Behaviour Using Latent Class Cluster Analysis

Sinziana I. Rasca[a,*], Karin Markvica[b], Benjamin Biesinger[b]

[a] Department of Engineering Sciences, Faculty of Engineering and Science, University of Agder, Jon Lilletuns vei 9, 4879, Grimstad, Norway
[b] AIT Austrian Institute of Technology GmbH, Center for Energy, Giefinggasse, 4, 1210 Vienna, Austria

ARTICLE INFO

Keywords:
Travel behaviour
Latent class cluster analysis
Personas
Transport
Urban networks

ABSTRACT

The present study proposes a new methodology that combines quantitative and qualitative data for the generation of representative personas for commuters. The profiles can be used to better understand their travel behaviour and mode choices. The research is based on the example of the region of Agder in Norway and aims to overcome the persona development shortcomings identified by previous researchers. Data from a regional travel behaviour survey (N= 1 849) is analysed using latent class cluster analysis (LCCA), and enriched with qualitative input from 32 interviews, and information provided by an expert panel. This results in a set of 20 representative persona profiles for the case study region. The proposed methodology is easily replicable in other urban networks and has the potential to provide insight into the mobility behaviour and needs of specific groups of people in order to adapt the transport services and encourage climate-friendly behaviour.

1. Introduction

Work trips are an essential part of the everyday life of adults. Such trips are performed regularly and routinely and take up a large part of our time budget for mobility (Ahmed and Stopher, 2014; Zahavi and Talvitie, 1980). As they contribute significantly to the volume of traffic, the mode choice for these trips plays a crucial role in reaching the 2030 European climate and energy framework targets (European Council, 2014). Directing a larger part of the commuter traffic to more sustainable transport modes, such as public transport (PT) or active mobility, is necessary for this purpose.

An essential prerequisite for the provision of a user-friendly transport system is the knowledge of user needs and requirements that determine an individual's transport mode choice (Filippi et al., 2013). This approach puts the user at the centre of the design process, in a similar way to the persona approach introduced by Cooper (1999) and defined as “a user-centered design (UCD) and human-computer interaction (HCI) technique that promotes immersion into end-users' needs” (Salminen et al., 2022). Personas are mostly employed in the fields of software development, healthcare, and higher education, as shown in the review of Salminen et al. (2022). Inspired by the potential of the persona approach to support the implementation of UCD in practice, we propose a methodology for producing persona profiles that would be ready to employ in the frame of travel behaviour research as stereotypical representations of transport user types. Employing personas in travel behaviour research has been attempted before (Mayas et al., 2014; Vallet et al., 2020), but multiple challenges in the design process of representative persona profiles, such as high cost, lack of objectivity and rigour, lack of scaling, non-representative data, risk of expiry, still need to be overcome (Chapman and Milham, 2006; McGinn and Kotamraju, 2008; Salminen et al., 2020).

* Correspondence
E-mail addresses: sinziana.rasca@uia.no, sinziana.rasca@gmail.com (S.I. Rasca).

https://doi.org/10.1016/j.multra.2023.100095
Received 30 November 2022; Received in revised form 14 February 2023; Accepted 5 March 2023

The present research aims to enable the generation of personas that can be incorporated into transport studies and passenger transport development work, based on a meaningful combination of different human factors emerging from the parallel analysis of quantitative and qualitative data sets. As a result, the scope of the present research is (a) to propose an improved approach for developing persona profiles, based on a mix of statistical analysis (latent class clustering) of sociodemographic characteristics and other factors that have proven influential on the travel behaviour characteristics, and on the analysis of qualitative data with the support of an expert panel, (b) to validate the approach through an expert panel assessment, and (c) to give guidance on persona development for other urban areas by highlighting the main steps in the creation process.

The proposed approach is applied to the geographical context of Agder region in southern Norway, which provides an interesting case study having key features typical of coastal Northern-European regions. The region, consisting of small cities and towns, struggles with reducing the car dependency of its inhabitants, mainly in the employees group.

To our knowledge, this is the first study that focuses on persona creation for employee travel behaviour purposes combining the following elements: (1) studying the context of networks of small cities and towns; (2) employing a latent class cluster analysis (LCCA) as a quantitative data analysis method; and (3) using readily available quantitative data for the generation of the base persona profiles and local experts plus a qualitative data set for the persona profile validation and enrichment. The process for the persona development should be easily replicable in different spatial contexts. As such, this paper contributes to the forefront of research within the field of travel behaviour.

The paper is organised as follows: the current state of the art in relation to persona creation and use in travel behaviour research, together with an overview of the use of LCCA in the same field is presented in Section 2. Section 3 presents the case study and Section 4 briefly reviews the data collection and analysis methods, the techniques employed for the persona generation and enrichment and the validation approach. Section 5 presents the results of the research in the form of the generated persona profiles, while Section 6 covers the discussion and further research directions emerging from the results of the study. Finally, Section 7 concludes our research with an overview of the most significant findings.

2. State of the art

The travel behaviour of employees is a very specific field as its manifestation is dependent on multiple factors, such as employment location, work schedule, or even household composition. Therefore the next sections concentrate on literature covering the current approaches in capturing mobility behaviour, and the advantages and challenges in using methods like personas and LCCA in travel behaviour research.

2.1. Current approaches in capturing mobility behaviour

Travel behaviour research is heavily reliant on the understanding of behavioural change. General behaviour change theories, such as the Theory of Planned Behaviour (TPB) (Ajzen, 1985) or the Theory of Interpersonal Behaviour (TIB) (Triandis, 1977) have been adopted by the transport research field and fine-tuned to the specificities of travel behaviour. TPB-based interventions have been proven effective by Steinmetz et al. (2016, p. 216) in a meta-analysis of 82 papers, with a mean effect size in behaviour change of .50, and with “effect sizes ranging from .14 to .68 for changes in antecedent variables (behavioural, normative, and control beliefs, attitude, subjective norm, perceived behavioural control, and intention)”. As the TPB “traces the causal links from beliefs, through attitudes and intentions, to actual behaviour” (Ajzen, 1985, p.11), it fits well with the concept and methodological approach proposed for the construct of persona profiles in the present study. Therefore, the TPB represents the starting point of our research.

According to Le Loo et al. (Le Loo et al., 2015, p.3), travel behaviour is influenced not only by psychological drivers, such as instrumental, symbolic, and affective motive, but also “by: (1) locational determinants; (2) socio-demographic characteristics; and (3) cultural attributes and perceptions”. We can observe a wide variety of research that focuses on the statistical analysis of variables falling in the three groups mentioned by Le Loo et al. (2015) and how they affect the travel behaviour of individuals in different settings (Balcombe et al., 2004; Chng et al., 2016; Ha et al., 2020; Ma and Ye, 2019). Their results generate a cross sectional image of how a selected variable directly impacts the use of specific transport modes in the case of the user group analysed.

Diverse statistical segmentation approaches have been elaborated in the field of passenger transport as early as 1979 (Sen and Benjamin, 1979), aiming to form homogenous groups according to different variables such as household composition or employment status, in order to enable more targeted research of travel behaviour formation (Chakrabarti and Joh, 2019; Chng et al., 2016). Some of the most popular methods employed for variable based segmentation in travel behaviour research are clustering (Fürst, 2014; Haustein and Jensen, 2018; Soto et al., 2021) and factor analysis (Beirão and Sarsfield Cabral, 2007; Li et al., 2020; Outwater et al., 2003). Classes of variables in mobility research segmentation mainly focus on mobility behaviour, spatial variables, socio-demographic and socio-economic variables, and psychographic factors (i.e., attitudes and values) (Haustein, 2012; Markvica et al., 2020). The segmentation approach is generally quantitative, being based on collected data sets of actual travel behaviour patterns (Haustein and Hunecke, 2013).

A more recent statistical segmentation approach is the latent class cluster analysis, or LCCA. This has been explored as a quantitative data segmentation solution for the passenger transport sector since the late 90’s (Goulias, 1999; Sasaki et al., 1999), proving to be “more efficient for considering individual taste heterogeneity” (Sasaki et al., 1999, p. 39). LCCA attempts to divide a given collection of input samples (called population) into a set of groups or classes based on the hypothesis that the class label is an unobserved categorical variable that divides the population into mutually exclusive and exhaustive latent classes (Lanza and Rhoades, 2013; Rafiq and McNally, 2021). Since its introduction to the transport sector, LCCA has been employed in a variety of travel behaviour studies that ranged from studying determinants of a certain behaviour (Beckman and Goulias, 2008; Kroesen, 2014) to building user typologies and profiles (Machado et al., 2018; Rafiq and McNally, 2021), but no examples of it being used for persona creation in this field could be identified. The TPB and LCCA have been jointly used by previous researchers in studying the travel behaviour and decision making of individuals (Fu, 2021).

The last decade has marked the emergence of a different approach in capturing mobility behaviour, one that includes a process of segmentation but moves beyond the rigidity of segmenting quantitative data (Vallet et al., 2020): the persona approach. Personas are stereotypes or archetypes of users (Cooper, 1999) and have initially been elaborated as a design tool for user-friendly software as they give insights on "the needs, goals and frustrations of users" (Brickey et al., 2012, p. 538). Working with personas is said to help moving from abstract goals to tangible assumptions (Vallet et al., 2020). Faily & Fleichas define personas as "behavioural specifications, embodying the salient characteristics of a class of stakeholders a design needs to serve" (2011, p. 2267).

The initial approach in persona creation was based on qualitative data and "look[ed] for people who clump together across multiple variables'' (Goodwin, 2002). More recent approaches for persona generation include quantitative, mixed and hybrid methods (Salminen et al., 2020). As not every user is covered, personas are a generalisation of potential behaviour types and do not claim completeness (Grudin and Pruitt, 2002; Salminen et al., 2020). Nevertheless, as contextual profiles are seen to "influence greatly a service's deployment and execution, since context-aware services should adapt to context and related updates" (Panagiotakis et al., 2005, p. 2014), the use of personas in travel behaviour research and public transport planning and operation shows excellent potential to support a better implementation of UCD approaches in this field.

2.2. Advantages and challenges of using personas

Personas allow researchers to go beyond segmentation (Salminen et al., 2020), offering an in-depth understanding of different personality, household, culture and environment related aspects that can influence a person's behaviour. Personas represent core tools in UCD, helping organisations communicate about the target user in a more empathic way (Salminen et al., 2020). According to Pruitt and Grudin (2003, p. 1), personas "provide a conduit for conveying a broad range of qualitative and quantitative data, and focus attention on aspects of design and use that other methods do not".

Even though personas can be created using a diversity of methods, they are still most often generated based on qualitative data (Brickey et al., 2012; Salminen et al., 2020). The traditional qualitative approach foresees that "each interviewee is mapped against the appropriate set of variables. In that way, a segmentation of the data into separate groupings based on these behavioural patterns is created, each presenting a different persona profile" (Laporte et al., 2012, p. 266).

The use of qualitative data alone is contested by diverse scholars in the persona creation process (Chapman and Milham, 2006; Salminen et al., 2020). Underlining this aspect, Gaiser et al. (2006, p. 521) argue that "In order to fulfill standards of a scientific method, personas can't be created arbitrarily. Personas have to be grounded in data, at best, both qualitative and quantitative data of surveys with the target audience.". Their statement is supported by several other researchers (Pruitt and Grudin, 2003; Salminen et al., 2022, 2020). The review of Salminen et al. (2020), that covers quantitative data-based persona development methodologies, highlights the rising popularity of the mixed method approach in persona generation, with 38.8% of the cases reviewed combining quantitative and qualitative methods.

The main challenges in creating personas based on qualitative data alone have been summarised by Salminen et al. (2020): high cost, lack of objectivity and rigour, lack of scaling, non-representative data, and the risk of expiry or becoming obsolete. Nevertheless, some criticism exists for both qualitative and quantitative persona creation. Salminen et al. (2020, p. 1-2) summarises the three main points in this respect: "the risk of personas being abstract and inaccurate", "simplifying complex human behaviours into simple archetypes that may be useful only to a degree", and "being "just" one method of user centric design while other methods can be better in some use cases". In addition, a reduced number of factors makes the results simpler to grasp whereas a large sample of factors provides a better fit with the data (Brickley 2012).

The mobility field benefits from regular access to both qualitative and quantitative data sets due to periodic user data collection done at national, regional and local level for the purpose of national statistics and also for understanding mobility trends. In this field, some examples for using persona profiles are: the development of new mobility services (Beyer and Müller, 2019; de Clerck et al., 2018), the improvement of communication with transport users (VDV 2011) as well as requirements, acceptance research for new mobility concepts (Kong et al., 2018), and scenario thinking for mobility interventions (Alonso-González et al., 2020; Vallet et al., 2020).

2.3. Tangible approaches in creating personas

Quantitative persona creation (QPC) has been found to address some of the shortages mentioned above in relation to using solely qualitative data. QPC is defined by Salminen et al. (2020, p. 2) as "using algorithmic methods to create accurate, representative, and up-to-date personas from numerical and textual data". Some of the advantages related to QPC mention: the scientific verifiability of personas and their increased credibility for stakeholders, as QPC use "real data"; the potential for statistical representativity, replicability and verifiability (Salminen et al., 2020); and "the availability and abundance of user data and the rapid development of data analysis algorithms" (Salminen et al., 2020, p. 2).

According to both previous research (Brickey et al., 2012) and the latest reviews of QPC studies (Salminen et al., 2022, 2020), clustering represents the most popular approach in persona creation, with more than one third of the examples reviewed by Salminen employing clustering methods such as k-means or hierarchical clustering. Automated statistical programs are widely employed for this purpose, (Brickey et al., 2012; Salminen et al., 2020), with a recent shift towards machine learning approaches in the last decade (Salminen et al., 2022, 2020). The cluster analysis is indeed suitable for large data sets but can provide clusters even for data lacking an underlying structure (Brickey et al., 2012; Hoerler et al., 2019).

The mentioned techniques alternate between the benefits and threats attached. Firstly, the selection of factors to be analysed in relation to the construct of the persona profile is currently not guided in any way, being left at the sole preference of the individuals who are designing the personas (Salminen et al., 2020). Secondly, as Gaiser et al. (2006) had already highlighted, factor analysis needs additional qualitative input to enrich the quantitative data, as key elements of the user behaviour can be overlooked otherwise.

Therefore, the present study combines the qualitative approach (interviews and expert panel) with a quantitative method (statistical analysis of survey data using LCCA) to create more reality-grounded persona profiles. At the same time, the approach proposes a solution that can address two other shortcomings identified in the classic persona creating process: the outdated data set and the high costs for data collection and analysis. Using travel survey data, which is often performed on a regular basis and available for low costs or even free in various countries and regions around the globe, the personas can be updated with minimum costs to ensure that they correctly represent the targeted population.

2.4. Latent Class Analysis and mobility behaviour research

The latent class structure theory was developed by Lazarsfeld in the 1950's (1950), offering multiple advantages against more traditional clustering techniques, such as K-means clustering (Rafiq and McNally, 2021). Porcu and Giambona (2016, p. 129) define LCCA as "a statistical method used to group individuals (cases, units) into classes (categories) of an unobserved (latent) variable on the basis of the responses made on a set of nominal, ordinal, or continuous observed variables". According to Molin et al. (2016, p. 15) "LCCA is a model-based approach that probabilistically assigns individuals to clusters and thus takes measurement error into account". The division is made using the hypothesis that the class label is an unobserved categorical variable that divides the population into mutually exclusive and exhaustive latent classes (Lanza and Rhoades, 2013; Vermunt and Magidson, 2002). In LCCA, the maximum-likelihood method is used to estimate the model parameters, so that "the identification of the latent mixture involves maximizing a log-likelihood function (just like in estimating structural equation models), which generates statistically consistent criterion (i.e., likelihood) for allocating individuals to the latent clusters" (Wang and Hanges, 2011, p. 26). LCCA can configure the classification and prediction of classes with a single maximum likelihood estimation algorithm simultaneously, also providing various goodness-of-fit measures, such as AIC (Akaike information criterion) or BIC (Bayesian information criterion), measures that are useful in determining the optimal number of classes (Rafiq and McNally, 2021). Such outputs are not available in K-means clustering models. Due to its model-based approach, LCCA allows both confirmatory and exploratory applications (Wang and Hanges, 2011).

LCCA has a broad range of applications, such as medical (Grant et al., 2020; Stout et al., 2018) and behaviour research (Mori et al., 2021; Wright et al., 2022), mainly focusing on the segmentation of user groups into representative classes. LCCA has also been widely used in travel behaviour research, covering topics such as: the classification of immigrants based on their commuting behaviour (Beckman and Goulias, 2008), modelling the behavioural determinants of travel behaviour (Kroesen, 2014), assessing the greenhouse gas impacts of different travel behaviour styles (Keskisaari et al., 2017), the interactions between the built environment, travel attitudes and travel behaviour (van de Coevering et al., 2018), the attitudes of individuals towards mobility as a service (Alonso-González et al., 2020), or the relation between habits and the commute mode decision process (Fu, 2021). Table 1 offers an overview of the most common LCCA uses in travel behaviour research based on the literature review performed for the purpose of the present study. As it can be observed in the table, the most popular software employed for performing LCCA analyses in the field of travel behaviour are: Latent Gold, Mplus, and R with the poLCCA (Polytomous Variable Latent Class Analysis) package.

Even though LCCA presents an excellent potential for supporting the development of personas, and clustering is the most popular segmentation method for quantitative data in persona generation approaches (Salminen et al., 2020), at the time of conducting the present research, no studies could be identified where LCCA has been employed to create persona profiles representative for the mobility behaviour of a population group. Despite its many applications in segmentation approaches, our literature review only identified one example of using LCCA in the process of creating persona profiles that focused on the analysis of artist types (Donze, 2011).

In the present research, the persona profiles will be constructed specifically for the field of transport research, using a mix of quantitative and qualitative data. The following two sections present the case study employed and the proposed methodological approach to building the persona profiles.

3. Case study

The region of Agder is situated in South-Eastern Norway with access to the North Sea coast. It covers a territory of 16,434 km^2 and has a population of just above 300 000 inhabitants, of which 112 000 inhabitants live in the municipality of Kristiansand. Approximately 80 percent of the population is concentrated on the coastal area, living predominantly in small cities and towns. The five largest urban areas in Agder are, in order of size, Kristiansand, Arendal, Lindesnes, Grimstad, and Vennesla, with the first four being located on the coast.

Public transport in Agder is almost exclusively bus-based, the services being offered by Agder Kollektivtrafikk AS (AKT). A limited rail network exists, but it is used mainly for long-distance trips as it does not connect the coastal municipalities and falls under the national rail service provider's jurisdiction, thus not having integrated fares with the rest of the public transport service. Confronted with low population densities, urban sprawl is an issue in the region and leads, together with easy access to motorised vehicles and free parking offered by employers, to a modal split of approximately 5% for PT (modal share is approximated from travel habit

Table 1
Overview of applications for LCCA in travel behaviour studies.

Author, year	*Title*	*LCCA use*	*Data set*	*Software employed*
Beckman and Goulias (2008)	Immigration, residential location, car ownership, and commuting behavior: a multivariate latent class analysis from California	Investigate the spatial, social, demographic, and economic determinants of immigrants' joint distribution among travel time, mode choice, and departure time for work.	Census form data, 2000	Not Available
Morin et al. (2010)	A Multifoci Person-Centered Perspective on Workplace Affective Commitment: A Latent Profile/Factor Mixture Analysis	Explore the usefulness of a person-centred perspective to the study of workplace affective commitment (WAC). Five distinct profiles of employees were hypothesised based on their levels of WAC directed toward seven foci (organisation, workgroup, supervisor, customers, job, work, and career).	Web-based questionnaire for Canadian employees, 2003	MPlus
Kroesen (2014)	Modeling the behavioral determinants of travel behavior: An application of latent transition analysis	Explore the notion that qualitative differences in travel behaviour patterns are substantively meaningful and therefore relevant from an explanatory point of view. Assess the effects of seven exogenous variables, including two important life events (i.e. moving house and changing jobs), on cluster membership and the transition probabilities.	5-year period mobility panel (survey and travel diary), 1984 - 1989	Latent Gold
Keskisaari et al. (2017)	Greenhouse gas impacts of different modality style classes using latent class travel behavior model	Analyse the interconnections between urban structure and socioeconomic, demographic and lifestyle variables and direct ground transport greenhouse gas (GHG) emissions. Identify and improve the understanding of the latent modality styles which guide people's everyday travel choices, and the resulting GHG implications.	Transport survey, 2012	SAS statistical program, utilizing the PROC LCCA extension
Van de Coevering et al. (2018)	Residential self-selection, reverse causality and residential dissonance. A latent class transition model of interactions between built environment, travel attitudes and travel behavior	Explore how people across different population groups adjust their residential environments and attitudes over time. Model interactions between the distance to railway stations and travel-mode related attitudes and the distance to shopping centres and the importance of satisfaction with these distances.	Internet questionnaire, 2005	Latent Gold 5.0
Machado et al. (2018)	Finding service quality improvement opportunities across different typologies of public transit customers	Provide a methodology of service quality evaluation based on PT customers behavioural theory and advanced market segmentation. Identify transit service improvement opportunities for specific customer typologies.	Online survey, 2014	Not Available
Alonso-González (2020)	Drivers and barriers in adopting Mobility as a Service (MaaS) – A latent class cluster analysis of attitudes	Identify factors relevant for MaaS adoption based on a survey with over thousand respondents in the Netherlands. We find five clusters in relation to individuals' inclinations to adopt MaaS.	Online survey, 2018	Latent GOLD 5.1
Yankholmes et al. (2021)	A latent class approach to examining migrant family travel behavior	Create empirically derived travel behaviour clusters of Western professional migrant families with and without children based on their motive to move, self-concept and how they construct a sense of home.	Online survey, 2016	Latent GOLD
Rafiq and McNally (2021)	Heterogeneity in Activity-travel Patterns of Public Transit Users: An Application of Latent Class Analysis	Analyse transit-based activity-travel patterns by classifying users via LCCA. Provide insights on the variations of activity-travel patterns and the associated market segments of transit users in the United States.	National Household Travel Survey, 2017	R Polytomous Variable Latent Class Analysis package
Fu (2021)	How habit moderates the commute mode decision process: integration of the theory of planned behavior and latent class choice model	Develop a comprehensive framework by integrating the theory of planned behaviour (TPB) and latent class choice model, aiming to understand how mode-use habits moderate commute mode choice	large-scale Household Travel Survey, Shaoxing County, China, 2012	Mplus 7

reports for the two main regions in Agder (Haugsbø et al., 2015a, 2015b)). This value is situated well below the national average of 11% (Grue et al., 2021). The local PT infrastructure is concentrated mainly on the municipality of Kristiansand, the rest of the municipalities having little to no local transport lines. At regional level, PT is ensured predominantly through regional bus lines on the coastal area, which connect the main municipalities.

The Norwegian National Travel Survey of 2018/19 found that Norwegians undertook an average of 3.26 trips per day, totalling 43.2 km and 71 minutes travelled, with an average trip length of 15,6 km and a duration of approximately 25 minutes (Grue et al., 2021). The survey also revealed that 85 percent of Norwegian households own at least one car even though 57 percent of the population reports having good or very good PT transport supply (Grue et al., 2021). This is very problematic in terms of greenhouse gas (GHG) emissions, as 17.8 percent of the total emissions released from the Norwegian territory in 2021 fall in the category of road

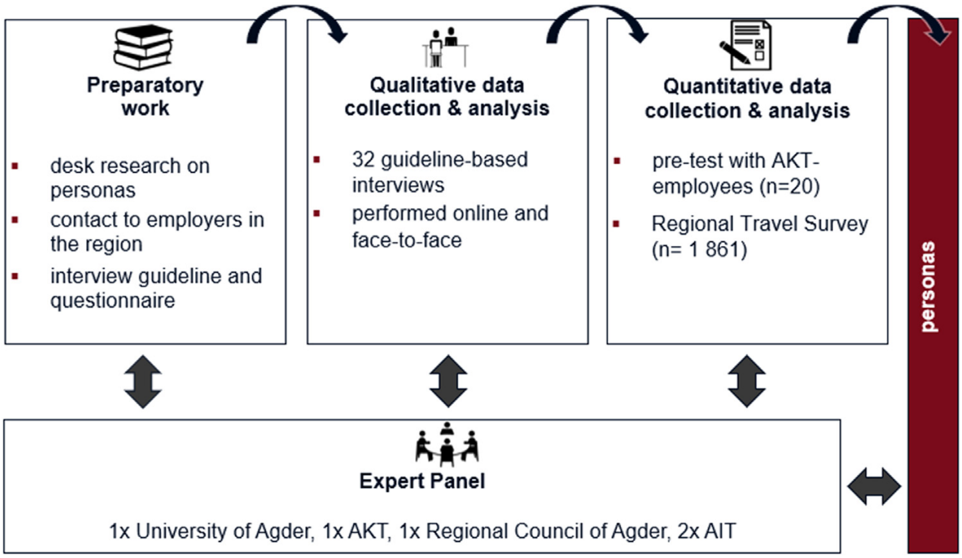

Figure 1. Overview of the steps for data collection and analysis.

traffic (SSB, 2021). The numbers extracted from the analysis of chronological travel surveys in Norway show the start of a downwards trend in the use of cars for the daily commute for Norwegians from 55% in 2013/2014 (Hjortol et al. 2014) to 53% in 2018/2019 (Grue et al, 2021). The latest national travel survey also found that "the car share for those having the longest education is somewhat lower than the average" (Grue et al., 2021, p. 4)

As it is assumed that most everyday journeys are repetitive (Thøgersen, 2006), it is likely that the mode choices made as part of daily routines are transferred to other types of trips. The so-called "habit generalization" leads to not choosing the mode of transport in a certain context, but rather transmitting a certain routine to different types of journeys (Garcia-Sierra et al., 2018).

Given the national priority of Norway to increase the use of PT (Norwegian Ministry of Transport, 2021), and that employees are one of the most car dependent groups in general, we chose to build representative personas for employees from a region in Norway.

To limit the focus of the research, the decision was taken to study only the daily commuting behaviour of employees, as a dominating, independent and highly mobile group, that has one of the highest contributions to road passenger traffic (Horner, 2004; Hunecke et al., 2007).

4. Data and Methods

In this section we introduce the data collection and the methods of analysis, together with the conceptual framework for the persona construct methodology. The methodology highlights the use of quantitative approaches, through the application of LCCA using Latent Gold 60F[1] as a statistical analysis software, and that of qualitative methods, in employing an expert panel for the structuring and validating a comprehensive set of persona profiles representative for employees aged 20 to 66 living and working in Agder, Norway. The resulting persona profiles are targeted at urban planners and the field of travel behaviour research, synthesising the dominant characteristics of representative employee groups in one geographical region.

4.1. Data collection and analysis

Based on the literature review in Section 2 covering the creation and use of personas, it was decided to use both qualitative and quantitative data in the persona profile creation process. Therefore, interviews were conducted to support the survey data already available. To make sure that all essential aspects of persona profile generation were covered, an expert panel was employed as a supporting function (see Figure 1).

4.1.1. Quantitative data collection through surveys

According to the review study of Salminen et al. (2020, p. 5), focusing on quantitative persona creation, "the most typical source for data collection is using surveys, with 55% of the articles reporting the use of surveys". Within our study, we employ quantitative data collected through a Regional Travel Survey (RS 2019) conducted between June and September 2019 in the frame of the OPTCORA project. The survey concentrated on employees living and working in the region of Agder, with an employment location

[1] https://www.statisticalinnovations.com/latent-gold-6-0/.

in close proximity to the route of regional bus line 100. Employers with over 100 employees (to preserve the anonymity of the respondents) were selected from the public Norwegian database of employers, Brønnøysundregistrene1F[2], and were sent an invitation letter describing the project, and purpose of the survey, with the request to distribute the survey among their employees. Several municipalities, private employers and the university in Agder chose to distribute the survey to their staff. The survey was available in Norwegian and English. No financial or material incentives were provided to either employers or respondents for distributing or answering the survey.

The survey consisted of 36 questions (multiple choice and open questions) focussing on five main thematic groups: demographic profile, work, commuting habits, public transport accessibility and satisfaction. 1 849 individuals provided complete answers to the survey.

4.1.2. Qualitative data collection through interviews

A series of 32 semi-structured interviews (see Table A1 in the Appendix for interview guidelines), representative for participatory approach as a methodology, was conducted to enhance the insights from the quantitative data collection. The interview questions were designed to verify the data collected through the survey. The interviews were performed online or face-to-face with employees of organisations that agreed to take part in our survey. The invitations to the interviews were distributed via the employers that participated in the survey to their staff. Direct contact to the full pool of employees was not an option due to privacy reasons. The data collection process was approved by the Norwegian Center for Research Data (Norsk senter for forskningsdata).

4.1.3. Role of the expert panel

The expert panel mentioned in Figure 1 was an element that also encompassed the participatory approach methodology. The panel was composed of five persons and it included: academic staff from the University of Agder working in the field of transport research; transport planners and decision makers from the public transport provider AKT and the Regional Council of Agder; and international transport researchers from the Center for Energy of the AIT Austrian Institute of Technology (one expert for modelling and agent-based simulations and one expert for mobility behaviour). The experts had different leading roles in the processes listed, connected to their principal expertise. For example, the desk research was mainly conducted by the three experts from academic and research institutions (UiA and AIT); the specific input on regional demographics and typical regional-transport behaviour were brought by the local partners (AKT; AFK, UiA); the potential connection with agent-based simulations was led by the AIT expert on the topic; the input on persona creation, travel behaviour, and qualitative data analysis was led by the mobility behaviour expert together with the UiA expert.

The experts had regular work meetings, where the topics relating to the current stage of the process were discussed. The leading experts would present their approach and findings, the results would be discussed in the panel, and a common panel decision would be taken on continuing or concluding the work. No major conflicts were registered, but differing opinions would be discussed, and a group agreement sought for moving forward.

The panel was given the role of a project advisory board. That means that it was informed on every step of the study and gave input and guidance in the different project phases. The expert panel gave input on the persona creation and validation processes based on the practical and academic knowledge of the experts in the field of transport.

4.2. Persona design and validation

Salminen et al. (2020, p. 2) synthesise the core four steps in a persona creation method, based on previous literature: “(a) data collection, (b) segmentation and grouping, (c) analysis of the qualitative and/or quantitative data, and (d) creating/writing persona profiles to present the user segments and their attributes as user archetypes”. Our proposed approach for the persona design process, presented in Figure 2, follows the same structure, employing an LCCA method for the segmentation phase and a participatory approach perspective (quantitative personas enriched and validated using qualitative data).

4.2.1. Starting point

The persona development, founded on the different sources mentioned, has been performed in accordance with previously gained knowledge from the “OptiMaaS - Optimized Mobility as a Service”2F[3] (OptiMaas) project. In OptiMaaS the persona profile relied on a structure incorporating eight field of characteristics, with three to nine elements each (OptiMaaS, 2020). The fields are: (1) socio-economic factors, (2) personality, (3) motivations, (4) expectations and values, (5) frustrations, (6) mobility choices, (7) influences and habits, (8) and frequently used information sources. Figure 3 presents the overview of the OptiMaaS persona structure, highlighting the eight fields of persona characteristics.

The structure of the OptiMaaS personas was found to be well suited to the purpose of the present research, but the methodology employed (based solely on desk research and qualitative analysis using focus expert panels) and the final selection of the personas themselves, were found lacking. Therefore, the present study uses a similar structure to the OptiMaaS persona, but not identical elements in each of the eight fields. This allows for the resulting personas to be more flexible and demographically representative for every age group.

[2] https://data.brreg.no/enhetsregisteret/oppslag/enheter?kommunenummer=0906&fraAntallAnsatte=100&sort=navn.norwegian,asc.

[3] https://www.optimaas.eu/.

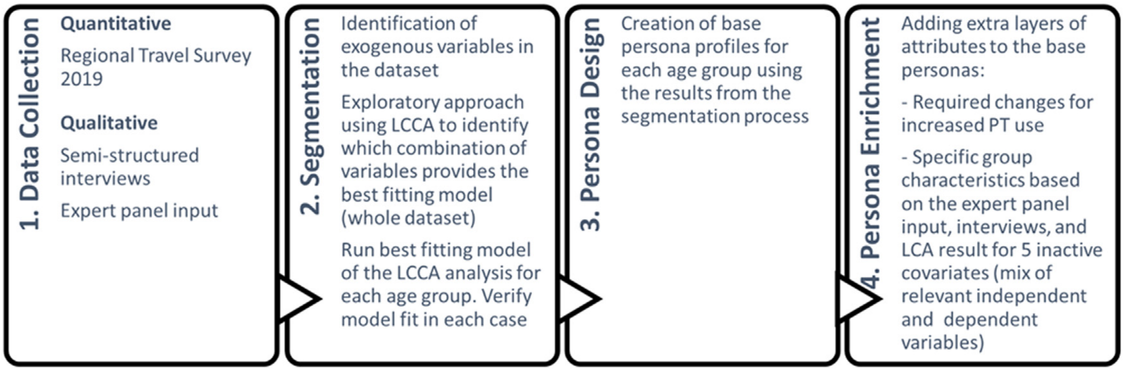

Figure 2. Steps of the persona design process.

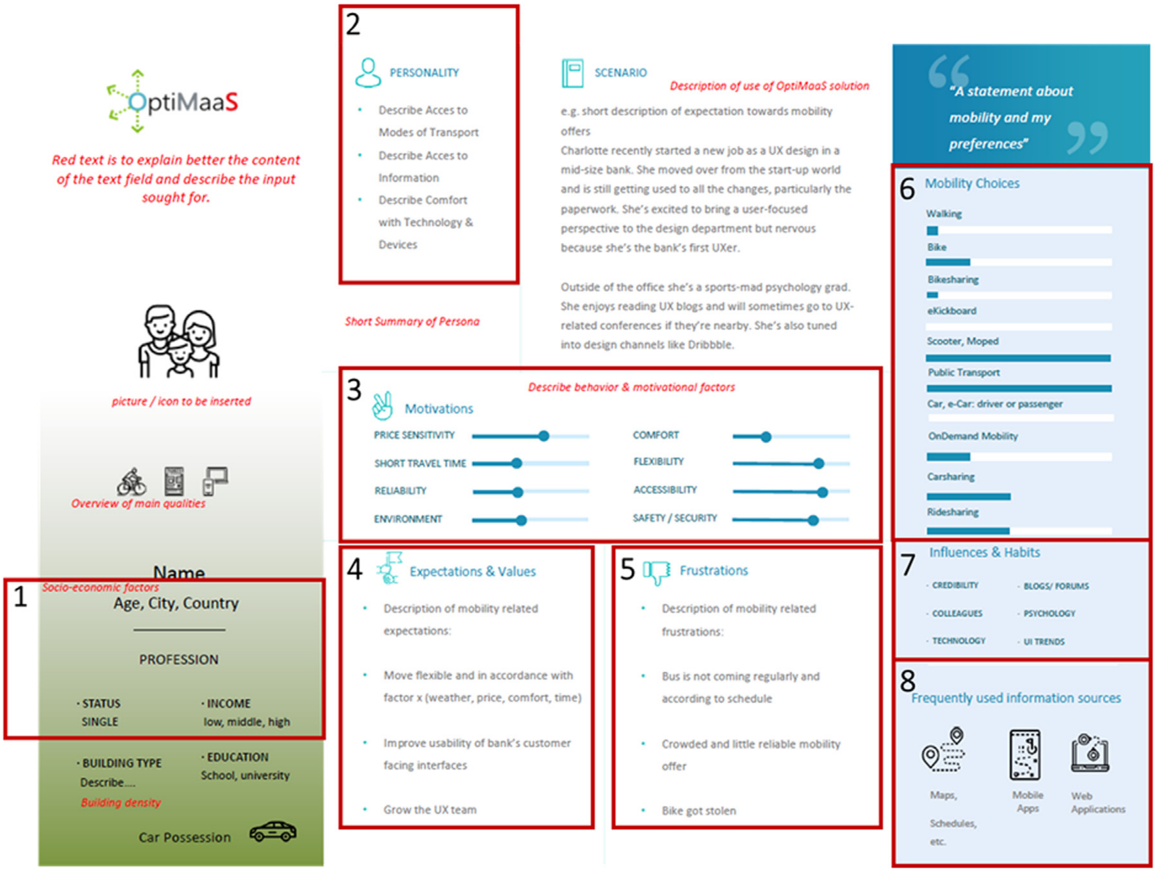

Figure 3. Overview of the persona profile structure in the OptiMaaS project.

As opposed to the OptiMaaS persona creation method, our approach combines both qualitative and quantitative data and methods. Furthermore, the use of a statistical analysis software for the segmentation gives more credibility to the process, the selection of different criteria for the persona profile having a mathematical motivation behind it. It also makes it easier and financially viable for the persona profiles to be updated regularly whenever travel behaviour data sets, such as national travel surveys, are available.

4.2.2. Segmentation process

The present research starts from the assumption that, with the presence of specific socio-demographic variables (age group, presence of children in the household, car ownership etc.), the majority of people will have travel-mode attitudes which are aligned to the characteristics of their socio-demographic group. For example, people having persons (mainly children) in care, would be more oriented towards car ownership and car use (Chakrabarti and Joh, 2019; Md Oakil et al., 2016), irrespective to the geographic and situational variables available to them (distance to work, frequency of PT etc.). In opposition, people with no persons in care would be more strongly influenced by geographic or situational variables present in their life in regards to their travel mode choice for the daily commute (Grue et al., 2021).

The purpose of the study is to identify representative groups with similar travel behaviours and attitudes towards various transport modes when the same geographical and/or situational factors influence their daily habits, within the targeted population. To reach the purpose, we use an LCCA approach to analyse data from the RS2019 and observe what kind of population segments tend to form in the studied population. Latent Gold 6 was employed for analysing the dataset and for the segmentation process.

For the purpose of conciseness, and due to the extensive number of studies dedicated to the LCCA method, we only present a synthetic overview of how LCCA was employed in the present analysis. We recommend the studies of Magidson et al. (2020), Porcu and Giambona (2016), and Rafiq and McNally (2021) for details about the LCCA methodology. Furthermore, for details about using the Latent Gold software, we recommend the work of Alonso-Gonzales et al. (2020) and Kroesen (2014), together with the software's technical guides (Vermunt and Magidson, 2013, 2008).

The LCCA model is built using exogenous, or independent variables, identified in the quantitative data set in relation to the mode choice for daily commuting. The exogenous variables mainly represent socio-demographic characteristics that are assumed to influence the membership of the profiles to different classes. An example could be that, based on the findings of the latest Norwegian national travel survey (Grue et al., 2021), males show a higher potential for belonging to a class with higher car use. Table 2 provides the descriptive statistics for the 11 exogenous variables identified in the available dataset, together with the abbreviations that will be used in the rest of the study to refer to them. As "the impact of constantly changing situational conditions like the weather, the goal of a trip, disruptions in travel services, etc. can only be understood meaningfully in a given situation and the impact is only valid for this specific situation" (Klöckner and Friedrichsmeier, 2011, p. 265), only variables that were not situational (e.g., frequency at the home bus-stop) or geographical (e.g., distance between home and work) could be considered exogenous (or independent) in relation to the travel mode choice.

The 11 exogenous variables can be grouped into five categories: socio-economic factors (*Gender, Age, Education, PersCare, CarOwn, License*), personality (*FitnessGoals*), motivations (*BusPrice*), work-related (*WorkSched, OffsiteWork*), and openness to PT (*RecommendPT*).

The indicators used for the LCCA analysis were socio-demographic indicators (*Gender, Age, Education, PersCare, CarOwn, License, WorkSched*) and a set of covariates related to a person's socio-demographic profile and travel behaviour (*OffsiteWork, FitnessGoals, BusPrice,RecommendPT*).

The first phase of performing the segmentation process using LCCA was exploratory. In this phase the covariates and indicators that should be eliminated in order to achieve a statistically significant model were identified. The statistical significance of the models, that contained between one and seven classes, was verified. To identify the best performing model, we reviewed the log likelihood p value ($p > .05$), the Bayesian information criterion (BIC), the Akaike information criterion (AIC) and AIC3 (LL) scores aiming for the smallest values. After each analysis, one covariate or indicator that presented a high bivariate residual (BVR) statistical value (significantly larger than four) was eliminated, and the analysis rerun. This technique is similar to the one employed by Porcu and Giambona (2016). It should be noted that no common agreement exists currently as to what combination of criteria is best fitted to determine the optimal number of classes in LCCA models (Nylund et al., 2007; Porcu and Giambona, 2016).

Once a statistically significant model was confirmed according to the aforementioned criteria for the entire dataset, the analysis could be continued with adding diverse inactive covariates and studying the distribution of the respondents in the different clusters.

4.2.3. Design of base profiles for the personas

The four clusters identified using the main dataset had to be tested for accuracy in relation to every age group (G1-G5). The distribution of the survey respondents into four classes with dominating features, as a result of the LCCA analysis, constituted the basis of structuring of the persona profiles. The exogenous variables that showed statistically significant results were further considered as core variables for the persona profile construct.

To avoid building persona profiles that include a strong age bias, as presented in the studies of Salminen et al. (2019) and Salminen et al. (2022) it was decided to run the same model for the five age groups studied, and only afterwards define the persona typologies, with the help of the expert panel input. The model fit was verified based on the same criteria and adjusted if necessary. The process followed the next steps for each age group:

- Test base profiles alignment with age group data by running the same LCCA model used in the initial segmentation phase.
- Identify cases where the base model is not statistically significant for different age groups and set variables with highest BVR values as inactive in the Latent Gold model (one at a time) until the model becomes statistically significant.
- Adjust the model, if necessary, by removing the variables with the highest BVR value.

The last step allowed for the detailed segmentation per age group, and for identifying similarities and differences between the formation of the four clusters in different age categories.

For the age group 20-30, rendering the two covariates inactive improved the model fit to a statistically significant level ($p=0.32$). For the age groups 31-40 and 61-66, the same action was successful, with a model fit value of $p=1$ and $p=0.97$ respectively. For the age groups 41-50 and 51-66, rendering only one of the two covariates inactive achieved statistically significant results. The best results in both cases were for rendering the *BusPrice* covariate inactive ($p=0.8$ and $p=0.89$, respectively).

As a result of the LCCA analysis, 20 clusters that formed the base of the persona profiles (four for each of the five age groups) were created using the selected set of exogenous variables. Once the four clusters for each age group were generated, it was visible that the

Table 2
Descriptive statistics of the 11 exogenous variables.

Question and abbreviation	***Answer categories***	***Descriptive statistics***					
		All data	***G1 20-30***	***G2 31-40***	***G3 41-50***	***G4 51-60***	***G5 61-66***
Gender	*Female*	57.9	58.3	59.8	58.1	57	57.9
(*Gender*)	*Male*	41.4	40.4	38.9	41.2	42.7	42.1
	Other	0.1	0	0	0.2	0	0
	I prefer not to say	0.6	1.3	1.3	0.5	0.4	0
Age	*Under 20*	0.4	x	x	x	x	x
(*Age*)	*20-30*	8.4	x	x	x	x	x
	31-40	20.7	x	x	x	x	x
	41-50	29.7	x	x	x	x	x
	51-60	28.8	x	x	x	x	x
	61-66	10.9	x	x	x	x	x
	67 and over	1.1	x	x	x	x	x
Education	*Higher education -long*	51.3	42.3	56.1	57.7	46.8	45.5
(*Education*)	*Higher education -short*	28.1	36.5	27.9	24.6	27.3	34.7
	Upper Secondary	18	20.5	14.1	15.1	22.9	15.8
	Lower Secondary	1.1	0	0.5	1.3	1.5	1.5
	I prefer not to say	1.5	0.6	1.3	1.3	1.5	2.5
Driving License	*Yes*	95.5	91.7	90.9	96.7	97.7	98.5
(*License*)	*No*	4.5	8.3	9.1	3.3	2.3	1.5
Persons in Care	*Yes*	48.8	13.5	74.4	75	30.6	6.4
(*PersCare*)	*Sometimes*	2.3	1.9	1.8	3.3	1.1	4
	No	49.3	84.6	23.8	21.7	68.2	89.6
Fitness Goals	*Not important*	9.6	9	12	9.3	8.8	8.4
(*FitnessGoals*)	*Very low imporance*	9.4	12.2	9.1	9.3	8.5	9.9
	Low imporance	12.5	19.9	12.8	12.8	10.7	10.4
	Moderately important	20.8	24.4	24.5	17.7	21.4	17.3
	Important	21.8	19.9	19.6	22.2	25.6	17.3
	Very important	17.8	12.8	15.4	20.8	17.1	20.8
	NA	8.1	1.9	6.5	8	7.9	15.8
Car ownership	*No car*	6.3	18.6	10.4	3.8	3.6	3
(*CarOwn*)	*1 car*	44.5	54.5	41	39.5	46.1	52
	2 cars	42	21.2	46	48.5	40.6	38.1
	3 cars or more	7.1	5.8	2.6	8.2	9.8	6.9
Work schedule type	Fixed	28.7	28.8	21.7	30.2	32.9	26.7
(*WorkSched*)	Flexible	50.4	45.5	56.7	50.3	48.1	49.5
	Shifts	11.6	23.2	14.1	10.2	10	5.4
	Other	1.4	0.6	1	1.3	1.1	2.5
	NA	8.1	1.9	6.5	8	7.9	15.8
Off-site for work purposes during working hours	Yes-daily	6.7	3.2	7	7.7	6.4	7.9
(*OffsiteWork*)	Sometimes	51.6	36.5	52	52.6	53.2	54.5
	No- never or rarely	41.7	60.3	41	39.7	40.4	37.6
Bus prices crucial for choosing transport mode	Yes	29.3	46.8	33.2	29.1	24.2	21.3
(*BusPrice*)	No	61.2	46.8	56.9	62.7	64.7	67.3
	Don't know	9.5	6.4	9.9	8.2	11.1	11.4
Likeliness to recommend PT to a colleague	*0 - Not at all likely*	23.8	19.9	21.9	23.5	23.7	22.8
(*RecommendPT*)	*1*	18.4	19.9	19.1	19.5	18.6	13.4
	2	11.6	17.3	11.5	11.8	10.3	10.9
	3	24.9	19.9	27.7	25.3	22.2	28.7
	4	4.8	5.8	4.2	4.2	4.9	5.4
	5	5.9	5.8	5.5	5.6	6.6	5.9
	6- Extremely likely	10.2	11.5	9.9	9.8	9.6	12.4
	NA	.3	0	0.3	0.2	0.6	0.5

*Age of the respondents corresponding to each of the five age groups..

corresponding cluster by number (one, two, three or four) in each of the age groups did not necessarily match for all five age groups in what concerns the profile features. Nevertheless, profile similarities could be observed between clusters with non-corresponding numbers in different age groups. Therefore, it was observed that four common persona base profiles could be achieved for all age groups if clusters with different numbers would be matched for this purpose. This aspect will be covered in more detail in the results section (see Table 7).

For the personas to be representative of the Agder demographics and credible in their profiling, a persona enrichment process was necessary. The enrichment process is described in the following subsection.

Table 3
General persona profiles for each age group agreed on by the expert panel.

Variable/Age group	*G1*	*G2*	*G3*	*G4*	*G5*
Age	20-30 yrs.	31-40 yrs.	41-50 yrs.	51-60 yrs.	61-66 yrs.
Income	+/++	++/+++	++/+++	++/+++	++/+++
Education	+/++	++/+++	++/+++	++/+++	++/+++
Child-care responsibilities	none	small children	children	older children	none
Mobile and tech-savvy	+++	+++	++	++	+
Flexibility	+++	+	++	++/+++	+++
Car dependency	+	+++	+++	+/++	+
Other	PT usage			interested in active commuting	interested in active commuting

+ low, ++ medium, +++ high

4.2.4. Persona enrichment

The enrichment process had a joint approach, combining the input of the expert panel and the outcomes of the LCCA. Before the LCCA analysis was performed, the expert panel had already proposed a set of simplified persona profiles based on the analysis of data from the 32 interviews and the experience of the panel members. Their proposal, presented in Table 3 under the Results section, was compared to the outcome of the LCCA model to identify both overlaps and inconsistencies.

The enrichment process started with adding other variables of interest, based on the OptiMaas persona structure, as inactive covariates to the LCCA model and analysing their distribution in the different clusters. After that, a new data layer covering “Specific characteristics” was introduced, based on the analysis of qualitative data set and on the input of the expert panel (corresponds to sections “Personality”, “Expectations & Values”, “Frustrations” and “Frequently used information sources” in the OptiMaaS persona structure).

5. Results

This section considers the results of both the quantitative and qualitative approaches employed in building the persona profiles for the employees in Agder. The results are presented in four separate subsections: Expert panel input, Quantitative data analysis, Final persona profiles, and Verification of persona profiles based on qualitative data. The first subsection includes the initial input of the expert panel on how the persona profiles would look like, the second presents the results of the LCCA analysis and the probability of an employee to belong to a certain class, the third gives the final results of the persona profile after combining the qualitative and quantitative input, while the last proposes a verification approach for the persona profiles.

5.1. Expert panel input

The expert panel input was provided in two stages: before and after the quantitative data analysis using the LCCA approach. The first stage consisted of a set of proposed persona profiles, based on the interview data and the expertise of the members of the panel. The second is the enrichment of the schematic persona profiles generated with the use of the LCCA approach. In this subsection we refer only to the input provided in the first stage, the input for the second stage is presented in Section 5.3.

Once the data collection was completed, but before the LCCA approach was implemented, the five experts worked together to create representative persona profiles for every age group used in the study. Table 3 below presents the outcomes of the profiles agreed on by the expert panel in the case of Agder. The profiles are not specific but cover the range of predominating traits for seven different essential factors in the persona profile construct. Specific mentions are put in the field “Other”. The experts used their knowledge of the area and users, together with information from the qualitative and quantitative data sets available for building the profiles.

It is visible that different age groups may have specific characteristics, such as childcare responsibilities (mainly present for G2-G3), which correlates strongly with the perceived level of flexibility for the members of these groups (very low for G2, medium for G3) and their car dependency (very high for G2-G3).

Another interesting aspect is the comfort with technology which is perceived to be declining together with a raise in age, starting with G3. At the same time, the two oldest age groups are perceived to have a raised interest in active commuting.

5.2. Quantitative data analysis

The LCCA model was initially constructed to include all 11 exogenous variables and was run for models with one to seven clusters, with all variables set as nominal. The first four LCCA analyses did not offer statistically significant results for any of the models studied. Table 4 presents the results for the fifth analysis, with one to seven clusters. The analysis used five indicators (*Gender, Education, PersCare, CarOwn, WorkSched*) and two active covariates (*OffsiteWork, BusPrice*), and presented statistically significant results for several models (two clusters or more). Therefore, its results were used for the further steps of the study.

To identify the best performing model in the fifth analysis, we reviewed the log likelihood p value ($p > .05$), the BIC, AIC and AIC3(LL) scores aiming for the smallest values. The BVR value was also reviewed, aiming for a result of Max. BVR < 4. The four-class

Table 4
Results of the LCCA analysis for seven variables, with models containing between one and seven clusters.

Clusters	*BIC (LL)*	*AIC (LL)*	*AIC3(LL)*	*Npar*	*df*	*p-value*	*Max.BVR*	*Class Error*
1	17478.848	17396.012	17411.012	15	1834	8.2e-15	24.9728	0.0000
2	17148.695	16955.411	16990.411	35	1814	0.23	25.8445	0.1182
3	17124.436	16820.704	16875.704	55	1794	0.97	3.3720	0.1739
4	17208.321	16794.141	16869.141	75	1774	1.00	3.4394	0.1933
5	17302.592	16777.964	16872.964	95	1754	1.00	2.6433	0.2388
6	17394.169	16759.093	16874.093	115	1734	1.00	1.8850	0.2238
7	17504.988	16759.434	16894.464	135	1714	1.00	3.7223	0.2501

Table 5
Item response probabilities of LCCA four classes (clusters) model and descriptive statistics for the selected variables (entire dataset).

Var. name	***Var. categ.***	***Descr. stat.***	***Response probability***			
			Cluster 1	***Cluster 2***	***Cluster 3***	***Cluster 4***
Class/Cluster size			42.7%	21.2%	19.2%	16.9%
Gender*	*Female*	58%	53.5%	**63.2%**	**71.6%**	46.7%
	Male	41%	46.2%	35.6%	28.4%	51.5%
Education	*Higher ed.- long*	51%	**72.4%**	35.6%	8.5%	**66.4%**
	Higher ed.- short	28%	23%	23.3%	46.8%	26%
	*Secondary**	19%	3.7%	38.9%	42.1%	6.4%
	Prefer not to say	1%	1%	2.2%	2.6%	1.5%
Pers. in care	*Yes*	48%	**69.7%**	47.7%	41.6%	3.42%
	No	49%	28.9%	49.9%	**56.2%**	**91.9%**
	Sometimes	2%	1.4%	2.4%	4.7%	4.7%
Car own.	*1 car*	45%	37.2%	40.9%	41.2%	**71.6%**
	2 cars	42%	**57.4%**	40.2%	46.1%	0.9%
	3+ cars	7%	5.4%	13.5%	8.1%	2.4%
	No car	6%	0%	5.3%	4.6%	25.4%
Work sched.	*Fixed*	29%	15.1%	**89.2%**	5.9%	13%
	Flexible	58%	**83.8%**	6.4%	35%	**85.3%**
	Shifts	12%	0.4%	0.2%	**59.1%**	11.6%
	other	1%	0.7%	4.2%	0%	1.6%
Covariate						
Off-site Work	*Yes- daily*	7%	8.1%	7.9%	5.8%	2.7%
	Sometimes	52%	**68.1%**	33.2%	25.1%	**63.1%**
	No - never	42%	23.8%	**58.9%**	**69.1%**	34.2%
Bus Price	*Yes*	29%	27.4%	23.5%	32.8%	37.2%
	No	61%	**64%**	**64.6%**	**57.4%**	**54.3%**
	Don't know	10%	8.6%	11.9%	9.8%	8.5%

*The response options for "Other" and "I prefer not to say" were excluded as they accounted for less than 1 percent of the total number of answers.
**Includes both lower and upper secondary education response options.

model presented in Table 4 showed the best results for the data (smallest AIC and AIC3, highest p-value, Max. BVR < 4). Once the four-class model was confirmed, the analysis could be continued with studying the distribution of the respondents in the different clusters, and even adding diverse inactive covariates to observe how their categories would be distributed among the four clusters.

The profile of the latent class membership was analysed for the four-class model based on the seven variables. The results, presenting the response probabilities for each item considered, are presented in Table 5. They include: (1) the class sizes based on unconditional class membership probabilities, and (2) the average values of the indicators and covariates conditional on class membership, or item response probability. The resulting class size shows that people have a high probability of belonging to the first class (42.7 percent respectively), and a medium probability to belong to the second, third or fourth class (21.2, 19.2, and 16.9 percent, respectively). The BVR results for all the variables considered (including covariates) had values smaller than four, revealing that all indicators have a significant influence on the transport mode choice. Thus, the indicators significantly discriminate between the clusters. Regarding the active covariates, both *OffsiteWork* and *BusPrice* have a significant influence on class membership for the four-class model, where the respondents are not sorted by age.

The model uncovered four classes in the complete dataset. We can observe the following predominant traits in the four cluster (please note that these descriptions do not cover the full heterogeneity within each class):

- Cluster 1: Car dependent, highly educated parents - higher education long, persons in care, 2 cars or more, flexible work schedule, sometimes off-site for work, bus price generally does not matter.
- Cluster 2: Car dependent women (both with persons in care and not) with less schedule flexibility - female, secondary education or higher education short, equal distribution for persons in care and no persons in care, with one car or more, predominantly fixed work schedule, rarely going off-site for work, bus price generally does not matter.

- Cluster 3: Car dependent women without persons in care- predominantly higher education and no persons in care, one or more cars, working in shifts, very rarely off-site for work, with little interest in the bus price.
- Cluster 4: Car dependent individuals (no persons in care) with high schedule flexibility - secondary education or short higher education, having persons in care, two cars or more per household, working on a fixed schedule, sometimes or daily off-site for work purposes, with little interest in the bus price.

The applicability of the profiles identified in the general dataset was then tested on the five age groups (G1-G5) individually. The outcomes of the clustering analysis performed with these variables, and verified for statistical significance, are presented in Table 6. The variables and covariates that were turned inactive for different age groups to reach statistical significance are marked in the table.

When running the four-cluster model on the age groups, it is to be expected that the respondents corresponding to cluster one across the five groups will not always have similar characteristics to each other. The results in Table 6 show that differences appear indeed between the five age groups. These differences were registered on account of, for example, different percentages of age group members working on a fixed schedule or in shifts, as is visible for G2 in Cluster 2. Therefore, adjusting the order of the clustering results is necessary, in a limited number of cases, to better group the results. This approach will be presented in the next section.

For each of the four personas and their respective age group clusters, the distribution of responses for inactive covariates is available in Table A 2 presented in the Appendix. Here we will shortly present an overview of the most interesting results related to the covariates, results that we believe reflect travel behaviour traits connected to the persona profiles. We consider a positive or negative variation of 20 percent between (1) the average number of respondents who chose the category of a covariate in their response, and (2) the number of respondents in the age group from a cluster who selected the same category, as a significant variation. The variations that can be related to sustainable travel behaviour are marked in green, while the ones that can be related to increased use of cars are marked in light red in Table A 2 in the Appendix. The most important behaviour traits, based on the inactive covariates, were summarised and introduced into the persona profiles that can be explored in the following section.

5.3. Persona formation based on LCCA results and qualitative input

Forming representative persona typologies that could then be differentiated by age group based on common traits necessitated a rearrangement of the clusters. This was performed by the expert panel, with the purpose of maximising the similarities between the clusters of respondents in the different age groups belonging to one typology of persona profile. The panel of experts compared the results of the LCCA analysis for the whole dataset, and for the individual age groups, with their initially drafted personal profiles (Table 3). Once the outliers were identified in the data from Table 6, the data could be rearranged to better fit together in more representative groups.

As can be seen in Table 7, for Persona type 1, only the cluster for the age group G4 had to be switched with the same age group in cluster number three. For the Persona type 2 the same approach was necessary for age group G2. Persona types 3 and 4 needed more rearrangement to generate a consistent profile. In the cases of Persona types 2 and 4, the respondents work in shifts or a fixed schedule. To simplify the profile format, the shift work has been allocated to Persona type 4, as it corresponds to the smallest cluster and the fixed work schedule was allocated to Persona type 2. This would more correctly represent the distribution of the descriptive statistics for the entire dataset.

The resulting persona typologies were further enriched with data from the set of inactive covariates (see Table A.2) and input from the expert group that was based on analysing the interview transcripts and their own experience. The qualitative data added some more depth to the specificities of each persona, which are summarised in the form of the field “Other special characteristics” in Table 8, where the final results were assembled in the form of the 20 persona profiles.

Analysing Table 8, we see that there are no identical personas across the age groups, the closest ones being Personas 2 and 3 for G2 and G3. Therefore, we have 20 different non-binary personas (four personas per age group), which can be distributed in situ according to the demographic data. Table 8 also presents how each persona profile relates to “Required changes to use public transport more”, information generated from analysing the results of the inactive covariates in the Appendix. Each age-group-related set of personas tends to have three to five common “required changes”. At the same time, every age group had up to four extra requirements for improvements. The most demanding group was G2, which corresponds to the group that has small children in care. Results show that all age groups demand primarily better frequency, shorter travel time and lower prices for the public transport, irrespective of the car ownership rate. Better timetables, in respect to the user’s needs, are requested mainly by persons who own two cars (in all age groups).

Persona type One (P1) is represented by individuals with predominantly higher education and flexible work schedules, for which generally the bus price is not crucial for the daily commute mode choice. Car ownership is generally high (one car or more for persons above the age of 50) and very high (persons aged 31 to 50). The very high car ownership is correlated with the presence of persons in care (G2 and G3). The only age group with a lower car ownership (one car or no car) and more interest in bus prices (bus prices considered crucial for 45 percent of the respondents) is G1, where respondents are aged 20 to 30 years old. The age groups with high and very high car ownership (G2-G5) also need to go off-site for work purposes sometimes or daily, while the respondents in G1, the group with the lowest car ownership, need to go off-site sometimes or almost never.

Persona type Two (P2) is represented by individuals with secondary and higher education, either predominantly (G1 and G2) or in equal distribution. These personas have work schedules that are either fixed or in shifts with few benefiting from flexible work schedules. Nevertheless, all persons in this group don’t need to go off-site for work purposes, with a small percentage needing to

Table 6
Results of LCCA four classes (clusters) models for each age group considered (G1-G5). All values represent percentages.

Var. name	Var. categ.	Descr. stat.	Cluster 1					Cluster 2					Cluster 3					Cluster 4				
		Age group	G1	G2	G3	G4	G5	G1	G2	G3	G4	G5	G1	G2	G3	G4	G5	G1	G2	G3	G4	G5
		Cluster size	40	61	57	34	53	23	15	25	25	30	21	13	11	25	9	17	11	7	16	8
	Female	58	**64**	**60**	**63**	50	46	51	**59**	**70**	**83**	**95**	**99**	**78**	**38**	42	21	3	37	7	**55**	41
Gender	Male	41	36	39	37	49	54	49	38	30	18	5	1	22	61	**58**	**79**	**89**	**61**	**87**	44	**59**
	Higher ed.- long	51	**95**	**67**	**81**	**60**	**63**	1	**89**	20	16	30	1	4	**57**	**78**	6	27	15	3	17	34
	Higher ed.- short	28	5	30	18	27	27	8	10	31	34	32	**98**	**57**	31	21	**66**	**73**	7	44	26	**57**
	Secondary	19	0	2	0	12	9	**88**	1	**47**	**49**	31	1	37	12	0	28	0	**79**	**45**	**50**	9
Education	Prefer not to say	1	0	2	0	1	1	3	0	3	0	7	0	2	0	0	0	0	0	8	7	0
	Yes	48	13	**92**	**89**	0	7	18	3	**68**	11	0	17	**76**	7	**70**	31	3	**75**	**90**	**65**	0
	No	49	**87**	8	11	**100**	**93**	**76**	**97**	29	**89**	**96**	**81**	24	**72**	26	**69**	**96**	14	7	35	**65**
Persons in care	Sometimes	2	0	1	0	0	0	7	0	3	0	4	2	0	21	5	0	0	12	3	0	35
	1 car	45	**55**	40	38	**59**	**58**	42	50	36	45	47	**61**	42	**64**	39	**57**	**61**	31	24	31	25
	2 cars	42	18	**55**	**55**	28	37	34	2	48	50	47	28	**58**	1	**55**	6	0	40	**68**	31	52
	3+ cars	7	3	3	7	6	3	14	0	13	0	0	6	0	5	6	37	0	8	8	39	23
Car ownership	No car	6	**23**	1	0	7	2	9	**48**	2	5	6	5	0	**30**	0	0	**39**	**22**	0	0	0
	Fixed	29	31	18	26	21	2	50	18	44	43	**60**	17	11	14	25	**82**	9	**61**	46	**55**	2
	Flexible	58	**69**	**80**	**73**	**75**	**98**	15	**80**	17	27	26	24	2	**85**	**75**	3	**72**	18	49	29	**68**
	Shifts	12	0	0	0	0	0	32	1	40	29	14	**60**	**87**	1	0	15	19	21	0	16	0
Work schedule	other	1	0	2	2	4	0	3	0	0	0	0	0	0	0	0	0	0	0	5	0	31
Covariate																						
	Yes- daily	7	2	8	7	3	6	5	5	5	3	7	3	8	1	12	13	3	5	33	11	17
	Sometimes	52	48	**58**	**64**	**72**	**67**	29	52	21	24	39	22	36	**58**	**68**	37	37	40	**66**	35	50
Off-site Work*	No - never	42	50	35	29	25	27	**66**	43	**74**	**73**	**54**	**75**	**56**	41	20	50	60	55	1	54	34
	Yes	29	45	32	29	25	25	47	42	28	23	16	44	32	37	27	22	**54**	32	24	22	16
	No	61	52	**61**	**62**	**65**	**66**	44	44	**64**	**68**	**73**	45	**56**	**59**	**60**	**61**	41	**54**	**68**	**66**	**68**
Bus Price**	Don't know	10	3	8	9	11	10	9	14	8	9	11	11	13	5	13	18	4	14	8	12	16

* Variable turned inactive for age groups G1, G2, G5.
** Variable turned inactive for all age groups.

Table 7
Rearrangement of age group clusters to form representative persona typologies.

Persona type 1					
Indicators	G1.C1	G2.C1	G3.C1	G4.C3	G5.C1
Gender*	Females dominate (>60%)			50/50*	
Education	Highly educated				
Persons in care (PiC)	No PiC	Persons in Care			No PiC
Car ownership	1 or no car	2 cars or more		1 car or more	
Work schedule	Flexible				
Covariates					
Offsite work purposes	No and sometimes	Sometimes (and daily)			
BusPrice Crucial	Yes 45% No 51.6%	No (>60%)			

Persona type 2					
Indicators	G1.C2	G2.C4	G3.C2	G4.C2	G5.C2
Gender*	50/50	Males domin. (>60%)	Females domin. (>60%)	Females domin. (>75%)	
Education	Secondary education		50/50 Highly educated and secondary		
Persons in care (PiC)	No PiC	Persons in Care		No PiC	
Car ownership	1 or no car		2 cars or more	1 car or more	
Work schedule	Fixed or shifts				Fixed or flexi.
Covariates					
Offsite work purposes	No	No and sometimes	No		
BusPrice Crucial	50/50 Yes and No	No (54-73%)			

Persona type 3					
Indicators	G1.C4	G2.C2	G3.C3	G4.C1	G5.C4
Gender*	Males domin. (>75%)	50/50	Males dominate (>60%)		50/50
Education	Highly educated				
Persons in care (PiC)	No PiC				
Car ownership	1 or no car			1 car or more	2 cars or more
Work schedule	Flexible				
Covariates					
Offsite work purposes	No and sometimes	Sometimes (and no)	No and sometimes	Sometimes (and daily)	
BusPrice Crucial	Yes 54% No 41%	50/50 Yes and No	No (>60%)		

Persona type 4					
Indicators	G1.C3	G2.C3	G3.C4	G4.C4	G5.C3
Gender*	Females dominate (>75%)		Males domin. (>75%)	50/50	Males domin. (>75%)
Education	Highly educated	60/40 Highly educated and secondary	45/55 Highly educated and secondary		Highly educated
Persons in care (PiC)	No PiC	Persons in Care			No PiC
Car ownership	1 car or more		2 cars or more	1 car or more	
Work schedule	Fixed or shifts		Fixed or flexi.	Fixed or shifts	
Covariates					
Offsite work purposes	No		Sometimes (and daily)	No and sometimes	
BusPrice Crucial	Yes 44% No 45%	No (54-68%)			

*50/50 – one of the groups dominates slightly (values between 50 and 60 percent)
Note: The colors used as cell backgrounds in the table signify the following: light green – response correlated by previous research findings with sustainable transport mode choices; light red – response correlated by previous research findings with increased car use; other colors (orange, purple, blue, yellow – swap between age groups of different persona typologies (same color marks groups that were swapped).

Table 8
Final persona profiles per age group.

	Persona	Education	Cars per household	Work schedule	Persons in Care (PiC)	Deliver PiC	Recommend PT*	PT price crucial	Required changes for increased PT use		Other special characteristics	
G1 age 20-30	P1	Uni. Long	🚗	Flexible	No	No	✓	Maybe	1.Lower prices for PT 2.Higher frequency 3.Shorter travel time 4.Direct connections	5.Environmental impact	1. Very comfortable with technology 2. Highly mobile 3. Open to new transport alternatives 4. Medium to low income	No children. More time flexibility. Would prefer better frequencies and longer timetables in the evenings. Car use for daily commute: lower than average. Preferred sustainable mobility alternatives: walking and PT.
	P2	Secondary		Fixed			✗ ✗			5. Better timetable		
	P3	Uni. Short	🚫🚗	Flexible			0	Yes				
	P4	Uni. Short	🚗 🚗	Shifts	Yes	Yes	✗ ✗	Maybe				Small children, highly dependent on parents. Sometimes need to get home on short notice. Car use for daily commute: higher than average, also due to work schedule. PT not considered a commuting alternative.
G2 age 31-40	P1	Uni. Long	🚗 🚗	Flexible	Yes	Yes	✗	No	1.Higher frequency 2.Lower prices for PT 3.Direct connection 4.Shorter travel time	5.Better timetable	1. Very comfortable with technology 2. Medium income	Small children, highly dependent on parents. Sometimes need to get home on short notice. High car dependency. Car is seen as a provider of freedom and flexibility. PT not considered a commuting alternative.
	P2	Secondary	🚗	Fixed			✗ ✗			5.Shorter waiting time at connection		
	P3	Uni. short		Flexible	No	No		Maybe		5.Better timetable		No children. More time flexibility. Would prefer better frequencies and longer timetables in the evenings Car use for daily commute: lower than average. Preferred sustainable mobility alternatives: walking, biking and PT.
	P4	Uni. Short	🚗 🚗	Shifts	Yes	No	✓	No		5.More incentives		Both small and older (more independent children). Slightly increased use of car, probably affected by work schedule, but also open to other transport means. Preferred modes: bus, bike and walk.
G3 age 41-50	P1	Uni. Long	🚗 🚗	Flexible	Yes	Yes	✗	No	1.Higher frequency 2.Lower prices for PT 3.Shorter travel time 4.Direct connections	5.Better timetable	1. Very comfortable with technology 2. Medium to high income	Both small and older (more independent children). Sometimes need to get home on short notice. High car dependency. Car is seen as a provider of freedom and flexibility. PT not considered a commuting alternative.
	P2	Secondary		Fixed		No	✗ ✗					
	P3	Uni. short	🚗	Flexible	No		✓			5.Environmental impact		No children. More time flexibility. Would prefer better frequencies. Lower car dependency. Preferred sustainable mobility alternatives: PT, walking, and biking.
	P4	Secondary	🚗 🚗	Shifts	Yes	Yes	✗ ✗			5.Better timetable		Older and more independent children. Sometimes need to get home on short notice. High car dependency as a necessity due to work schedule. Car is seen as a provider of freedom and flexibility. PT not considered a commuting alternative.

(continued on next page)

Table 8 (*continued*)

G4 age 51-60	P1	Uni. Long	🚗🚗	Flexible	Yes	No	X	No	1. Lower prices for PT 2. Higher frequency 3. Shorter travel time 4. Direct connections 5. Better timetable		1. Comfortable with technology 2. Medium to high income	Older and more independent children. Sometimes need to get home on short notice. Medium car dependency. Car is seen as a provider of freedom and flexibility. Preferred sustainable mobility alternative: biking.
	P2	Secondary		Fixed	No		XX					No children. More time flexibility. Medium car dependency. Would like to have better frequencies for PT. Preferred sustainable mobility alternative: bike, bus and walk.
	P3	Uni. short	🚗	Flexible			X					
	P4	Secondary		Shifts	Yes		XX					Older and more independent children. Sometimes need to get home on short notice. High car dependency as a necessity due to work schedule. Car is seen as a provider of freedom and flexibility. Preferred sustainable mobility alternative: biking.
G5 age 61-66	P1	Uni. Long	🚗	Flexible	No	No	V	No	1.Higher frequency 2.Lower prices for PT	3.Shorter travel time 4.Direct connections 5.Better timetable	1. Not that comfortable with technology 2. Medium to high income 3. Medium car dependency for daily commuting	No children in the household. Would like to have better frequencies for PT. Preferred sustainable mobility alternative: bike, walk and bus.
	P2	Secondary		Fixed			XX			3. Direct connections 4. Shorter travel time 5. Better timetable		
	P3	Uni. short	🚗🚗	Flexible						3.Better timetable 4.Shorter travel time 5.Direct connections		No children in the household. Medium car dependency as a remaining habit- life without at least one car is not easy to imagine. Car is seen as a provider of freedom and flexibility. Preferred sustainable mobility alternative: bike, walk and bus.
	P4	Uni. short	🚗	Shifts	Yes					3.Direct connections 4.Shorter travel time 5.Better timetable		Independent persons in care in the household. Car needed due to work schedule- life without at least one car is not easy to imagine. Car is seen as a provider of freedom and flexibility. Would prefer better frequencies for PT. Preferred sustainable mobility alternative: bike and walk.

* Willingness to recommend PT: XX-very low, X-low, 0-neutral, V-high.

go off-site sometimes. Similar to Persona One, the bus price is not crucial for the daily commute mode choice in most cases, except for G1 where bus prices are considered crucial for 50 percent of the respondents. Car ownership is generally high and very high for persons above the age of 40 (2 cars or more for persons aged 41 to 50, and one car or more for persons above the age of 50). The very high car ownership is correlated with the presence of persons in care (G3). Two age groups have a lower car ownership (one car or no car): G1 and G2.

Persona type Three (P3) is represented by individuals with predominantly higher education, and flexible work schedules, similar to Persona One. Unlike Persona One, in this case all age groups are characterised by not having persons in care. For the Persona

Three type of respondents the bus price is not crucial for the daily commute mode choice in the cases of G3, G4 and G5, but we see an increase in interest for the bus price for both G1 and G2. In this case the car ownership is generally lower than for Personas One and Two, G1 to G3 having only one car or no car per household, G4 one car or more, and only G5 presenting a very high car ownership rate (two cars or more). A correlation between the two youngest age groups, lower car ownership (one car or no car) and more interest in bus prices (bus prices considered crucial for minimum 50 percent of the respondents) was observed. The two oldest age groups with high and very high car ownership also need to go off-site for work purposes sometimes or daily, while the respondents in the other age groups have a much lower need to go off-site for work purposes during work hours.

Persona type Four (P4) is represented by individuals with predominantly fixed or shift work schedules, similar to Persona Two. Also similar to Personas Two and One, we observe the presence of persons in care for three out of the five age groups (G2-G4), and a reduced interest in the bus price as a crucial factor for daily commute choices for respondents above the age of 30. When car ownership is considered, Persona Four shows what is probably the highest car ownership rate compared to the other three personas, with G3 having two or more cars per household, and the rest of the age groups one or more cars. We observe a correlation between the car ownership rate and being off-site for work purposes, with G3 having to go off-site more than the other four groups. In what concerns the aspect of education, Persona Four stands out with a mix of education levels for the five age groups, ranging from highly educated (G1 and G5) to a proportionate mix of secondary and higher education for G2, G3 and G4.

The generated persona profiles are schematic, but they can easily be further developed for application purposes, by allocating a gender, a profile picture, and a story to each of them, based on the core information developed with the help of the present method. The initial generation of the persona typologies presented in Table 7 can in fact support the allocation of genders to the personas.

The resulting profiles bring a further contribution to the UCD perspective through the summary of the Required changes for PT use for each of the personas. These required changes can offer to the PT planners, PTAs and local authorities targeted insights into what changes are valued by the potential users and how specific population groups would respond to diverse improvements of the PT service. For example, the reduction of ticket fares is a top priority only for G1 and G4, coming in second for the rest of the age groups. Moreover, it is obvious that a higher frequency of PT would serve all employee types.

6. Discussion and further research

The present research gives insights into how to approach a persona construction process in the field of passenger transport, using a blend of quantitative and qualitative data, and a mixed method approach built around an LCCA model. It also presents a set of required changes in the PT system from the user perspective. In this section we will discuss the results achieved and highlight potential avenues for further research.

The representativity of personas for dominating demographic groups has been questioned by researchers (Chapman and Milham, 2006). The results achieved for the present case study show strong potential for overcoming this deficiency by combining the construction of profiles for age groups with the use of LCCA in generating the base profiles of the personas, especially when we compare the schematic profiles designed by the expert panel with the results achieved through the LCCA analysis. This approach ensures a better chance of success in correctly representing a demographic group, being replicable to other population segments. The approach could be easily extrapolated to the whole population, given that the necessary data sets are available. Further research is recommended to explore how the proposed persona generation approach can be employed with publicly available datasets. It would also be interesting to explore how applying the methodology to longitudinal data sets from different years affects the profile structure of the personas and their distribution in different age groups.

The data collection for the persona profiles design has been identified as a hindering factor as it is time and cost intensive (Drego et al., 2010). The risk of personas becoming outdated (An et al., 2018) is also a reality in a rapidly changing society. Our approach in using data from a regional travel survey, that has similar questions to many national travel surveys, shows that one can use regularly occurring surveys, with publicly available data, to construct and regularly update persona profiles for a specific field. If the method is proven effective it could be used in any geographical setting provided that a sufficiently large set of representative quantitative data exists for the respective area, thus overcoming the aforementioned challenges. The limits imposed by the use of data sets not specifically designed for persona creation need to be noted, with data such as income level missing in the present research for that reason. Nevertheless, the common practice of using readily available data for creating persona profiles shows that there is value in creating and using personas despite their lack of complete accuracy (Salminen et al., 2022).

The travel behaviour of a persona can change dramatically based on variables that are independent of the psychological profile (Le Loo et al., 2015). It is impossible to create a set of personas that can cover all these situations. This shows the necessity of separating the persona profile data, mainly composed of independent variables (age, gender, persons in care) from the geographic and situation variables that should be allocated to the personas based on statistical data. The method proposed in this paper does not focus on geographical variables. Nevertheless, the potential impact of geographical and situational variables on the travel mode choice of the identified personas should be explored, especially if these profiles are to be further used in defining agents in agent-based simulations. Moreover, if personas are to be used in PT forecasting, it will be necessary to find a modality to code their profiles into numeric format. This is another point that should be taken into a further research approach.

The resulting personas have a variety of potential uses, ranging from transport planning approaches, where the planners and network designers can maintain better focus on the user groups and needs by employing the persona profiles in the planning processes, to the improvement of agent profiles in agent-based simulations. Furthermore, they can be used in communicating with policy makers and even the user population, humanising the data and information being transmitted. It would be interesting to further explore the added value of using personas in these types of scenarios, as compared to the regular operations now in place.

7. Conclusions

Improving the overall transport system in a region and providing user-friendly transport services poses several challenges. User-centred design in transport planning and operations requires in-depth knowledge on the travel behaviour and needs of the persons living in the area. To achieve this type of knowledge, it is necessary to bring together both statistical analyses of large data sets collected for the target population group in order to identify the main parameters, and qualitative data that offers further insight into why certain behaviours may occur. Creating personas offers many benefits to the transport providers and planners as shared attitudes and needs can be identified and tailor-made concepts set in place addressing each group targeted. The approach is much easier to grasp than purely statistical evaluations and can also be better communicated when different disciplines work together.

The LCCA analysis used in the current study for Agder identified five core variables as essential in defining the user groups: *Gender, Education, PersCare, CarOwn, WorkSched*. Therefore, the conclusion is that personas should represent both people with higher and lower levels of education which have access to diverse personal motorization levels (no car, one car, and two cars or more per household), have diverse household compositions reflective of the age groups, and work on various schedules. Interestingly, the cost of public transport services turned out to be of low importance for the case study area in Norway.

We believe that the methodology proposed in the present research for generating personas can be easily reproduced in other settings, with the use of publicly available travel survey data, a small set of interviews providing additional background information and a local panel of experts in the transport field setting a focus. It must be noted that publicly conducted surveys are often standardised and purposefully collected to be representative for all population groups of a region or a country, regularly updated to reflect changes in the travel behaviour of a certain population, and the data generated by them is often available at no cost. Furthermore, the proposed approach brings a new perspective into the credibility of persona generation and builds a solid bridge between population segmentation approaches where LCCA is employed and QPC.

We can thus conclude that the proposed hybrid methodology for persona creation is easily scalable to national level and applicable to various population groups and geographical contexts (for which data can be procured), has a reduced cost and can be regularly updated (subject to the availability of public data from regular travel behaviour surveys). Nevertheless, the limitations imposed by the data collection on the results should be considered, where the data is not collected in the frame of regular national travel surveys. This brings up the topic of objectivity and rigour, where we can conclude that the methodology progresses towards the creation of more accurate personas through the use of both LCCA, and the input of an expert group, thus aiming to overcome the shortcomings of single method approaches as described in Section 2.3.

As the passenger transport field is undergoing critical changes, with the introduction of new transport models such as Mobility as a Service and breakthrough technologies in the shape of self-driving vehicles, it is crucial to better understand the user needs and behaviours in shaping the future of passenger mobility. Therefore, we see a wide range of potential applications and contributions for the outcomes of our research, such as:

- Offering a quick, and easy to relate to, radiography of the target population group from the perspective of demographic profiles associated to the travel behaviour potential.
- Advancing of persona use in the field of public transport planning and operations with the purpose of better understanding and addressing the needs of the customers when undertaking mobility improvement actions.
- Providing a user-centred discussion base, representing the major demographic groups, for policy, planning, and operation in the mobility field.
- Supporting PT providers in customising their offer and providing better services to the representative groups identified.
- Using persona profiles to elaborate targeted communication and information strategies and campaigns that support the modal shift towards more climate-friendly transport modes.
- Bridging the gap between persona creation and the profile definition of agents in agent-based simulations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is part of the scientific projects “OPTCORA - Optimization of the Public Transport in the Coastal Region of Agder” (Grant from the Regional Research Fund Agder, Project number: 309046) and “OptiMaaS - Optimized Mobility as a Service” (Grant from the Austrian Research Promotion Agency, Funding frame: JPI Urban Europe - Making Cities Work, Project number: 872541). Special thanks to Agder Kollektivtrafikk AS, Matthias Prandtstetter and the AIT Austrian Institute of Technology - Center for Energy, Bynett Sør and William Fagerheim, Klimapartnere, the County Governor and the County Council of Agder, the municipalities of Arendal, Grimstad and Kristiansand for the support offered. The survey was distributed with the support of Klimapartnere, Bynett Sør and Sorlandet Teknologipark.

Appendix

Table A1. Travel behaviour interview guidelines (OPTCORA project, Agder, Norway, 2019).

Section	Question / Guidelines
Self-description (10-12 min)	Gender
	Age
	Education
	Type of work
	Do you have a driver's license? Which type?
	Do you own a car/have a car available for use? Which type of car?
	Work schedule (shifts, starting time, ending time)
	Place of residence (approx. distance to work)
	Living situation (alone, with partner, with partners, with children/ other persons in care)
	Income (low, medium, high)
	Did you relocate (move to where you are living now) in the past 3 to 5 Years?
	Tell about the relocation situation: When did they move, from where, why
	Personal interests
	Environmental concerns
	What are your usual activities in a regular weekday?
	What is your time budget for each?
	Is it easy to take decisions affecting your everyday life, or do you have to plan things in detail?
	What do you usually think about when you choose a mode of transport?
Mobility in general (6-8 min)	What's the role/importance of mobility in your everyday life?
	Could you tell me the first things that come in mind when I say:
	transport commute car public transport bike
	How do you organize your mobility? What means of transport do you use?
	How do you inform yourself/ plan the trip?
	What are the most important things for you when you chose your type of transport?
	Do you always use the same travel modes, or do you decide anew every time (I bike today, I drive tomorrow)? According to what criteria?
	If applicable: How did you organize your mobility at the previous place of residence, what means of transport, what information, how flexible, differences of everyday/leisure, etc.?
Mobility trends detailed questions (if not yet addressed) (4-6 min)	How do you feel about cars?
	For which trips do you use the car in everyday life? Why do you do them by car?
	Is it easy and cheap to park when you use the car?
	If you don't have a car, what are the reasons? (car-free, car-less)
	What are the best three things about using/having a car?
	What are the worst three things about using/having a car?
	How do you feel about sharing a car ride? Would you prefer to be the driver or passenger?
	Do you think some policy measures could be effective in convincing you/people to give up the car?
	Would you ever give up your car? What would have to happen for you to do that?
Public Transport (8-10 min)	What was the last time you used PT?
	Monthly pass?
	What everyday trips do you do by PT? Why do you do them by PT?
	If you go to work by PT, how long does that take? (door to door)
	And if you would mix modes?
	How do you feel about the Public transport options you have for your daily trips?
	Do you have PT stops closeby to your home/work?
	How frequent are the buses there?
	Does the PT stop have a shelter/ bench/ bike park/ car park/ lights/ electronic display?
	Would PT be a more economical option for you to commute?
	What are the best three things about using PT?
	What are the worst three things about using PT?
	Did you see any improvements in the PT offer in Agder in the last 3 years?
	If you don't use PT, what would make you give it a try?
Walking / Biking (2-4 min)	Do you care about fitness goals?
	Do you cycle/walk in your everyday life or for leisure?
	Why do you/don't you? What is good/bad about it?
Spatial influence (2-3 min)	What services do you have on your route?
	How strong do the availability of such services influence your trips (mobility path)?
	What is generally important to you in terms of mobility offers (offer, information, reliability, flexibility, comfort, security, privacy (*Let talk openly, then divide into external factors and internal factors)*
	How do you see an ideal development of mobility in the coming years?
	How would you envision a more sustainable everyday life in the future?

Table A2. Results by persona types and age groups for all inactive variables.

Inactive Covariate	Answer Categories	Descr stat	Persona type 1					Persona type 2					Persona type 3					Persona type 4				
			G1. C1	G2. C1	G3. C1	G4. C3	G5. C1	G1. C2	G2. C4	G3. C2	G4. C2	G5. C2	G1. C4	G2. C2	G3. C3	G4. C1	G5. C4	G1. C3	G2. C3	G3. C4	G4. C4	G5. C3
Driving License	Yes	96	92	98	99	99	99	89	95	96	98	96	92	81	85	97	100	94	65	97	98	100
	No	4	8	2	1	1	1	11	5	4	2	4	8	19	15	3	0	6	35	3	2	0
Deliver Pers In Care	Yes	17	10	58	21	12	1	16	47	11	1	0	0	42	5	0	0	16	2	25	6	0
	Sometimes	11	0	18	21	11	1	3	20	19	1	2	3	23	10	0	18	0	0	26	5	10
	No	72	90	24	58	77	98	82	33	68	98	98	97	32	85	100	82	84	98	49	89	91
Fitness Goals (1-Not important, 6-Very important)	1	10	7	12	9	10	6	6	14	9	10	12	18	14	9	9	1	9	8	12	6	15
	2	9	9	11	7	9	11	16	2	9	7	8	17	9	20	8	3	9	10	15	11	12
	3	13	20	12	14	10	7	19	12	9	10	13	22	18	17	10	6	18	12	10	15	22
	4	21	19	24	17	22	19	32	26	19	22	15	18	20	18	22	30	30	27	19	19	1
	5	22	28	19	22	24	15	12	24	26	26	19	14	17	16	25	12	17	22	20	29	29
	6	18	13	13	19	16	16	14	22	27	21	25	9	22	16	16	38	14	15	23	16	20
	NA	8	3	9	13	9	25	0	0	1	5	6	1	0	4	11	11	2	6	0	3	0
Use of PT Apps	AKT Billett	16	21	20	14	15	10	2	1:	16	18	8	19	14	24	16	22	20	21	20	11	6
	AKT Reise	10	16	8	10	11	10	6	7	10	10	4	20	7	13	14	3	11	14	6	7	0
	Both apps	25	45	33	30	20	17	41	21	18	15	12	45	26	31	21	9	26	41	15	13	18
	Don't use	41	14	30	34	44	37	31	54	55	51	70	14	52	28	38	56	40	19	59	66	75
	NA	8	3	9	13	9	25	0	0	1	5	6	1	0	4	11	11	2	6	0	3	0
Recommend PT (1- Not likely at all, 7- Most likely)	1	24	12	19	21	23	13	19	30	31	34	31	20	34	14	21	25	36	17	35	36	48
	2	18	14	19	18	18	12	25	26	22	22	15	16	18	15	14	16	29	15	29	26	12
	3	12	19	13	12	10	12	27	15	13	10	10	11	10	10	11	8	9	4	12	9	7
	4	25	24	30	29	27	33	17	17	17	18	23	21	21	32	26	42	13	31	16	14	14
	5	5	8	4	4	8	7	6	6	4	3	4	3	3	4	5	4	3	4	4	3	1
	6	6	9	5	7	8	8	0	1	4	5	3	12	1	5	7	3	0	14	1	6	5
	7	10	14	9	9	7	14	6	4	9	7	14	16	14	20	16	2	9	16	4	5	8
PT Service Rating	Very good	11	12	9	8	12	16	0	6	7	13	12	11	7	29	16	10	3	23	6	8	13
	Good	23	28	22	28	21	21	26	14	19	21	20	32	20	26	25	6	18	27	16	19	20
	Average	22	19	31	22	19	19	21	26	22	20	15	24	24	22	19	24	17	22	19	21	8
	Bad	12	15	12	11	11	13	21	10	13	12	13	18	5	7	10	16	17	10	15	19	13
	Very bad	19	12	15	19	25	18	23	22	24	18	22	7	30	10	15	18	31	6	31	22	16
	Don't know	13	12	10	12	11	12	8	21	16	16	17	7	14	7	13	27	14	13	14	12	30
T Mode CAR drive	Every day	43	25	52	44	42	26	59	69	55	48	45	20	50	24	29	34	61	18	61	58	37
	Often	15	5	12	20	18	21	11	20	13	15	14	4	13	12	13	21	14	6	28	13	23
	Occasionally	8	5	8	12	10	10	8	2	5	8	11	6	2	2	12	14	3	6	0	4	7
	Rarely	11	18	11	12	12	14	9	3	9	7	8	30	7	18	16	9	9	9	6	8	8
	Never	23	46	18	12	18	28	12	5	18	23	22	40	29	43	31	21	13	62	5	17	25
T Mode BUS	Every day	7	16	4	4	4	10	15	4	7	6	9	26	16	19	9	2	7	15	0	3	0
	Often	6	10	7	6	4	3	6	1	4	5	1	15	4	9	5	4	6	13	3	3	0
	Occasionally	8	6	7	9	14	10	6	3	4	7	6	5	5	14	13	3	0	14	6	5	2
	Rarely	12	21	13	13	16	13	9	11	8	8	6	20	11	10	16	6	11	8	12	8	15
	Never	67	47	69	68	62	63	64	81	78	74	78	34	65	48	57	84	76	51	79	80	83
T Mode BIKE	Every day	9	10	11	7	9	15	0	1	6	9	4	10	8	8	9	17	0	17	3	11	12
	Often	11	13	9	13	14	19	9	6	6	10	11	8	9	9	18	18	3	18	3	7	16
	Occasionally	10	8	7	12	12	15	4	5	11	12	10	8	7	11	11	11	12	10	14	8	10
	Rarely	9	5	12	12	8	4	11	13	6	6	2	2	8	14	8	2	5	7	18	12	1
	Never	61	65	61	57	56	47	77	76	71	62	74	71	69	58	54	51	79	48	62	62	61

(continued on next page)

(*continued*)

T Mode WALK	*Every day*	8	22	3	5	7	7	3	2	5	6	10	17	11	15	13	1	3	27	0	3	11
	Often	5	17	4	5	5	7	3	2	6	3	9	11	0	7	7	1	5	11	1	2	1
	Occasionally	6	6	4	6	8	9	3	4	6	5	5	10	7	8	7	17	11	10	1	4	18
	Rarely	7	8	7	9	4	5	3	16	4	6	4	5	9	11	6	12	2	16	6	4	1
	Never	74	48	83	75	76	72	88	76	80	80	73	57	72	59	67	68	78	37	93	87	69
*Better Prices	*Yes*	50	68	47	54	45	42	70	59	51	50	3	63	54	54	4	53	63	55	39	52	55
	No	50	32	53	46	55	58	30	41	49	50	67	37	46	46	56	47	37	45	61	48	45
*More Frequency	*Yes*	50	48	50	57	43	50	61	53	47	42	43	67	57	50	44	57	56	52	50	44	58
	No	50	52	50	43	57	50	39	47	53	58	57	33	43	50	56	43	44	48	50	56	42
*Shorter Trav. Time	*Yes*	38	52	46	38	40	32	46	44	36	38	34	35	38	34	30	30	60	30	32	36	21
	No	62	48	54	62	60	68	54	56	64	62	66	65	62	66	70	70	40	70	68	64	79
*More Direct Connections	*Yes*	33	34	32	29	33	28	38	45	41	38	37	40	38	25	25	25	55	35	29	35	27
	No	67	66	68	71	67	72	62	55	59	62	63	60	62	75	75	75	45	65	71	65	73
*Shorter Connection	*Yes*	21	25	18	16	19	10	29	40	25	26	16	32	28	25	16	22	17	18	20	24	27
	No	80	75	82	84	81	90	71	60	75	74	84	68	72	75	84	78	83	82	80	76	73
*BetterTimetable	*Yes*	26	28	24	21	23	20	38	34	32	33	25	34	31	14	21	39	43	17	30	29	22
	No	74	72	76	79	77	80	62	66	68	67	75	66	69	86	80	61	57	83	70	71	78
*Environmental Impact	*Yes*	19	3	20	18	20	17	15	22	22	17	18	24	17	26	21	14	22	15	14	14	8%
	No	81	69	80	82	80	83	85	78	78	83	82	76	83	74	79	86	78	85	86	86	92
*More Incentives	*Yes*	17	26	18	14	19	12	25	18	14	18	14	13	21	21	18	14	24	22	17	13	14
	No	83	74	82	86	81	88	75	82	86	82	86	87	79	79	82	86	76	78	83	87	86
*Could Work On Commute	*Yes*	12	22	17	12	8	6	18	19	7	13	15	25	14	14	7	21	13	8	1	8	2
	No	88	78	83	88	92	94	82	81	93	87	85	75	86	86	93	79	87	92	83	92	98
*Parking More Expensive	*Yes*	11	17	16	11	11	6	12	16	6	11	12	13	11	17	9	9	10	9	4	8	8
	No	89	83	84	89	89	94	89	84	94	89	88	87	89	83	91	91	90	91	96	92	92
*Better Park and Ride	*Yes*	10	13	14	11	7	9	3	20	10	9	9	16	12	7	7	17	9	12	12	6	1
	No	90	87	86	89	93	91	97	80	90	91	91	84	88	93	93	83	91	88	88	94	99
*Wish Tolls Higher	*Yes*	9	20	16	8	8	7	6	12	7	7	5	15	17	7	4	8	6	9	6	6	1
	No	91	80	84	92	92	93	94	88	93	93	95	85	83	93	96	92	95	91	94	94	99

Note: The colors used as cell backgrounds in the table signify a positive or negative variation of 20 percent between (1) the average number of respondents who chose the category of a covariate in their response, and (2) the number of respondents in the age group from a cluster who selected the same category. The light green color– response correlated by previous research findings with sustainable transport mode choices. The light red color– response correlated by previous research findings with increased car use; other colors (orange, purple, blue, yellow – swap between age groups of different persona typologies (same color marks groups that were swapped).

References

Ahmed, A., Stopher, P., 2014. Seventy Minutes Plus or Minus 10 — A Review of Travel Time Budget Studies. Transport Reviews 34, 607–625. doi:10.1080/01441647.2014.946460.

Ajzen, I., 1985. From Intentions to Actions: A Theory of Planned Behavior. Action Control, pp. 11–39.

Alonso-González, M.J., Hoogendoorn-Lanser, S., van Oort, N., Cats, O., Hoogendoorn, S., 2020. Drivers and barriers in adopting Mobility as a Service (MaaS) – A latent class cluster analysis of attitudes.Transp. Res Part A Policy Pract 132, 378–401.

An, J., Kwak, H., Jung, S., Salminen, J., Admad, M., Jansen, B., 2018. Imaginary People Representing Real Numbers.ACM Transactions on the Web (TWEB), p. 12.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. (Transportation Research Laboratory Report TRL593). Transportation Research Laboratory, London, UK.

Beckman, J.D., Goulias, K.G., 2008. Immigration, residential location, car ownership, and commuting behavior: A multivariate latent class analysis from California. Transportation (Amst) 35, 655–671.

Beirão, G., Sarsfield Cabral, J.A., 2007. Understanding attitudes towards public transport and private car: A qualitative study. Transp Policy (Oxf) 14, 478–489.

Beyer, S., Müller, A., 2019. Evaluation of Persona-Based User Scenarios in Vehicle Development. In: Advances in Intelligent Systems and Computing. Springer Verlag, pp. 750–756.

Brickey, J., Walczak, S., Burgess, T., 2012. Comparing semi-automated clustering methods for persona development.IEEE. Transactions on Software Engineering 38, 537–546.

Chakrabarti, S., Joh, K., 2019. The effect of parenthood on travel behavior: Evidence from the California household travel survey.Transp. Res Part A Policy Pract 120, 101–115.

Chapman, C.N., Milham, R.P., 2006. The personas' new clothes: Methodological and practical arguments against a popular method. In: Proceedings of the Human Factors and Ergonomics Society, pp. 634–636.

Chng, S., White, M., Abraham, C., Skippon, S., 2016. Commuting and wellbeing in London: The roles of commute mode and local public transport connectivity.Prev. Med (Baltim) 88, 182–188.

de Clerck, Q., van Lier, T., Messagie, M., Macharis, C., van Mierlo, J., Vanhaverbeke, L., 2018. Total Cost for Society: A persona-based analysis of electric and conventional vehicles.Transp Res D Transp. Environ 64, 90–110.

van de Coevering, P., Maat, K., van Wee, B., 2018. Residential self-selection, reverse causality and residential dissonance. A latent class transition model of interactions between the built environment, travel attitudes and travel behavior. Transp Res Part A Policy Pract 118, 466–479.

Cooper, A., 1999. The inmates are running the asylum. Sams.

Donze, P.L., 2011. Popular music, identity, and sexualization: A latent class analysis of artist. types.Poetics 39, 44–63.

Drego, V.L., Dorsey, M., Burns, M., Catino, S., 2010. The ROI Of Personas [Report]. Forrester Research.

European Council, 2014. Conclusions of the European Council Meeting (EUCO 169/14). Brussels.

Faily, S., Flechais, I., 2011. Persona cases: A technique for grounding personas. Conference on Human Factors in Computing Systems - Proceedings 2267–2270.

Filippi, F., Fusco, G., Nanni, U., 2013. User Empowerment and Advanced Public. Transport Solutions.Procedia Soc Behav Sci 87, 3–17.

Fu, X., 2021. How habit moderates the commute mode decision process: integration of the theory of planned behavior and latent class choice model.Transportation. (Amst) 48, 2681–2707.

Fürst, E., 2014. Making the way to the university environmentally sustainable: A segmentation approach.Transp Res D Transp. Environ 31, 1–12.

Gaiser, B., Panke, S., Arnold, P., 2006. Community design-the Personas approach. In: E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Association for the Advancement of Computing in Education (AACE), pp. 520–525.

Garcia-Sierra, M., Miralles-Guasch, C., Martínez-Melo, M., Marquet, O., 2018. Empirical analysis of travellers' routine choice of means of transport in Barcelona. Spain.Transp Res Part F Traffic Psychol Behav 55, 365–379.

Goodwin, K., 2002. Getting from research to personas: harnessing the power of data [WWW Document]. URL https://www.cooper.com/journal/2008/05/getting_from_research_to_perso/ (accessed 3.24.20).

Goulias, K.G., 1999. Longitudinal analysis of activity and travel pattern dynamics using generalized mixed Markov latent class models.Transportation. Research Part B: Methodological 33, 535–558.

Grant, R.W., McCloskey, J., Hatfield, M., Uratsu, C., Ralston, J.D., Bayliss, E., Kennedy, C.J., 2020. Use of Latent Class Analysis and k-Means Clustering to Identify Complex Patient Profiles. JAMA Netw Open 3 e2029068–e2029068.

Grudin, J., Pruitt, J., 2002. Personas, Participatory Design and Product Development: An Infrastructure for Engagement.

Grue, B., Landa-Mata, I., Flotve, B.L., (TOI), I. of T.E., Administration, N.P.R., Directorate, N.R., SF, B., AS, N.V., AS, A., 2021. The Norwegian national travel survey 2018/19 – key results. TØI Report 198p.

Ha, J., Lee, S., Ko, J., 2020. Unraveling the impact of travel time, cost, and transit burdens on commute mode choice for different income and age groups.Transp Res Part A Policy. Pract 141, 147–166.

Haugsbø, S.M., Ellis, I.O., Johansson, M., 2015a. Reisevaner i Kristiansandsregionen 2013/14.

Haugsbø, S.M., Ellis, I.O., Johansson, M., 2015b. Reisevaner i Arendalsregionen 2013/14.

Haustein, S., 2012. Transportation (Amst) 39, 1079–1103.

Haustein, S., Hunecke, M., 2013. Curr Opin Environ Sustain 5, 197–204.

Haustein, S., Jensen, A.F., 2018. Factors of electric vehicle adoption: A comparison of conventional and electric car users based on an extended theory of planned behavior. Int. J. Sustain. Transp. 12, 484–496. doi:10.1080/15568318.2017.1398790.

Hoerler, R., Haerri, F., Hoppe, M., 2019. New solutions in sustainable commuting-the attitudes and experience of European stakeholders and experts in Switzerland. Soc. Sci 8.

Horner, M.W., 2004. Spatial dimensions of urban commuting: A review of major issues and their implications for future geographic research.Professional. Geographer 56, 160–173.

Hunecke, M., Haustein, S., Grischkat, S., Böhler, S., 2007. Psychological, sociodemographic, and infrastructural factors as determinants of ecological impact caused by mobility behavior.J. Environ Psychol 27, 277–292.

Keskisaari, V., Ottelin, J., Heinonen, J., 2017. Greenhouse gas impacts of different modality style classes using latent class travel behavior model. J Transp. Geogr 65, 155–164.

Klöckner, C.A., Friedrichsmeier, T., 2011. A multi-level approach to travel mode choice - How person characteristics and situation specific aspects determine car use in a student sample.Transp. Res Part F Traffic Psychol Behav 14, 261–277.

Kong, P., Cornet, H., Frenkler, F., 2018. Personas and emotional design for public service robots: A case study with autonomous vehicles in public transportation. In: Proceedings - 2018 International Conference on Cyberworlds. CW 2018. Institute of Electrical and Electronics Engineers Inc., pp. 284–287.

Kroesen, M., 2014. Modeling the behavioral determinants of travel behavior: An application of latent transition analysis. Transp. Res Part A Policy Pract 65, 56–67.

Lanza, S.T., Rhoades, B.L., 2013. Latent class analysis: an alternative perspective on subgroup analysis in prevention. and treatment. Prevention science 14, 157–168.

Laporte, L., Slegers, K., de Grooff, D., 2012. Using correspondence analysis to monitor the persona segmentation process. In: NordiCHI 2012: Making Sense Through Design - Proceedings of the 7th Nordic Conference on Human-Computer Interaction. ACM Press, New York, New York, USA, pp. 265–274.

Lazarsfeld, P.F., 1950. The logical and mathematical foundation of latent structure analysis. Studies in Social Psychology in World War II Vol. IV: Measurement and Prediction, pp. 362–412.

Le Loo, L.Y., Corcoran, J., Mateo-Babiano, D., Zahnow, R., 2015. Transport mode choice in South East Asia: Investigating the relationship between transport users' perception and travel behaviour in Johor Bahru, Malaysia. J Transp Geogr 46, 99–111.

Li, X., Tang, J., Hu, X., Wang, W., 2020. Assessing intercity multimodal choice behavior in a Touristy City: A factor analysis. J. Transp. Geogr 86, 102776.

Ma, L., Ye, R., 2019. Does daily commuting behavior matter to employee productivity? J Transp Geogr 76, 130–141.

Machado, J.L., de Oña, R., Diez-Mesa, F., de Oña, J., 2018. Finding service quality improvement opportunities across different typologies of public transit customers 14, 761–783. doi:10.1080/23249935.2018.1434257.

Magidson, J., Vermunt, Jeroen K., Madura, John P., 2020. Latent Class Analysis, 2020th ed SAGE Publications Limited. SAGE Publications Ltd, Thousand Oaks, CA, USA.

Markvica, K., Millonig, A., Haufe, N., Leodolter, M., 2020. Promoting active mobility behavior by addressing information target groups: The case of Austria. J Transp Geogr 83, 102664.

Mayas, C., Hörold, S., Wienken, T., Krömker, H., 2014. One day in the life of a persona-a framework to define mobility agendas. In: Advances in Human Aspects of Transportation: Part III. AHFE Conference, pp. 211–218.

McGinn, J., Kotamraju, N., 2008. Data-driven persona development. In: Conference on Human Factors in Computing Systems - Proceedings. ACM Press, New York, New York, USA, pp. 1521–1524.

Md Oakil, A.T., Manting, D., Nijland, H., 2016. Dynamics in car ownership: The role of entry into parenthood.European. Journal of Transport and Infrastructure Research 16, 661–673.

Molin, E., Mokhtarian, P., Kroesen, M., 2016. Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers.Transp. Res Part A Policy Pract 83, 14–29.

Mori, C., Choi, H.J., Temple, J.R., Madigan, S., 2021. Patterns of sexting and sexual behaviors in youth: A Latent Class Analysis. J Adolesc 88, 97–106.

Morin, A.J.S., Morizot, J., Bastien Boudrias, J.-S., Madore, I., 2010. A Multifoci Person-Centered Perspective on Workplace Affective Commitment: A Latent Profile/Factor Mixture Analysis. Organ Res Methods 14, 58–90.

Norwegian Ministry of Transport, 2021.
Nylund, K.L., Asparouhov, T., Muthén, B.O., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study.Struct. Equ Modeling 14, 535–569.
OptiMaaS, 2020. How stereotypes help implementing MaaS – OptiMaaS [WWW Document]. URL https://www.optimaas.eu/?p=973 (accessed 9.4.22).
Outwater, M.L., Castleberry, S., Shiftan, Y., Ben-Akiva, M., Zhou, Y.S., Kuppam, A., 2003. Attitudinal Market Segmentation Approach to Mode Choice and Ridership Forecasting. Structural Equation Modeling https://doi.org/10.3141/1854-04, 32–42.
Panagiotakis, S., Koutsopoulou, M., Alonistioti, A., Thomopoulos, S., 2005. Context sensitive user profiling for customised service provision in mobile environments.IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. PIMRC 3, 2014–2018.
Porcu, M., Giambona, F., 2016. Introduction to Latent Class Analysis With Applications 37, 129–158. doi:10.1177/0272431616648452.
Pruitt, J., Grudin, J., 2003. Personas: Practice and theory.Proceedings of the 2003 Conference on Designing for User Experiences. DUX '03.
Rafiq, R., McNally, M.G., 2021. Heterogeneity in Activity-travel Patterns of Public Transit Users: An Application of Latent Class Analysis.Transp. Res Part A Policy Pract 152, 1–18.
Salminen, J., Guan, K., Jung, S.G., Chowdhury, S.A., Jansen, B.J., 2020. A Literature Review of Quantitative Persona Creation. Conference on Human Factors in Computing Systems - Proceedings.
Salminen, J., Guan, K.W., Jung, S.-G., Jansen, B.J., 2022. Use Cases for Design Personas: A Systematic Review and New Frontiers; Use Cases for Design Personas: A Systematic Review and New Frontiers. CHI Conference on Human Factors in Computing Systems 1–21.
Sasaki, K., Morikawa, T., Kawakami, S., 1999. A DISCRETE CHOICE MODEL WITH TASTE HETEROGENEITY USING SP, RP AND ATTRIBUTE IMPORTANCE RATINGS.World Transport Research 1 to 4 A1-A14\n2625 Vols.
Sen, L., Benjamin, J., 1979. Travel Behavior and Market Segmentation of Low and Middle-income Residents of Richmond. The Office, Virginia.
Soto, J.J., Cantillo, V., Arellana, J., 2021. Market segmentation for incentivising sustainable transport policies.Transp Res D Transp. Environ 99, 103013.
SSB, 2021. Emissions to air [WWW Document]. URL https://www.ssb.no/en/natur-og-miljo/forurensning-og-klima/statistikk/utslipp-til-luft (accessed 8.15.22).
Steinmetz, H., Knappstein, M., Ajzen, I., Schmidt, P., Kabst, R., 2016. How effective are behavior change interventions based on the theory of planned behavior?: A three-level meta analysis. Zeitschrift fur Psychologie /Journal of Psychology 224, 216–233.
Stout, J.E., Wu, Y., Ho, C.S., Pettit, A.C., Feng, P.J., Katz, D.J., Ghosh, S., Venkatappa, T., Luo, R., 2018. Evaluating latent tuberculosis infection diagnostics using latent class. analysis.Thorax 73, 1062–1070.
Thøgersen, J., 2006. Understanding repetitive travel mode choices in a stable context: A panel study approach. Transp Res Part A Policy Pract 40, 621–638.
Triandis, H.C., 1977. Interpersonal behavior. Brooks/Cole Publishing Company.
Vallet, F., Puchinger, J., Millonig, A., Lamé, G., Nicolaï, I., 2020. Tangible futures: combining scenario thinking and personas—A pilot study. on urban mobility.Futures 117, 102513.
Vermunt, J.K., Magidson, J., 2002. Applied latent class analysis 11, 60.
Vermunt, J.K., Magidson, J., 2008. UPGRADE MANUAL FOR LATENT GOLD BASIC, ADVANCED/SYNTAX, AND CHOICE VERSION 6.0 1.
Vermunt, J.K., Magidson, J., 2013. Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax. Statistical Innovations Inc, Belmont, MA.
Wang, M., Hanges, P.J., 2011. Latent class procedures: Applications to organizational research. Organ Res Methods 14, 24–31.
Wright, L., Steptoe, A., Fancourt, D., 2022. Patterns of compliance with COVID-19 preventive behaviours: a latent class analysis of 20 000 UK adults. J Epidemiol Community Health 76, 247–253.
Yankholmes, A., McKercher, B., Williams, N.L., 2021. A latent class approach to examining migrant family travel behavior. Tour. Manag 87, 104387.
Zahavi, Y., Talvitie, A., 1980. Regularities in travel time and money expenditures.