Technical University of Munich

Department of Civil, Geo and Environmental Engineering

Chair of Transportation Systems Engineering

MASTER'S THESIS

User Acceptance and Adoption of Driverless Shuttles: A Longitudinal Study

Author:
Xiaoyan Kang

Supervised by:
Dr. Jia Guo
Prof. Dr. Constantinos Antoniou

17th June 2021

Abstract

Operating driverless shuttles as a feeder transport service has been advocated as the most suitable way to integrate shared automated vehicles into the existing transportation system. To achieve this vision, understanding what factors influence the public's acceptance and adoption of the shuttle service and how these factors change over time are critical. Many studies based on pilot demonstrations have explored the factors influencing user acceptance of driverless shuttle with real-life ridden experience. However, these studies were restricted to predicting the adoption of the driverless shuttle by utilizing cross-sectional data and thus with no data to explore whether attitudes would change over time. To fill the research gap, a longitudinal survey is conducted. Using the three-wave panel data, the present study focuses on the real-world riding experiences of the driverless shuttles operated in a mixed traffic environment on public roads in Stockholm.

To better understand the longitudinal growth pattern of service evaluations that influence the overall satisfaction, a latent curve analysis is adopted. Results show that respondents evaluated speed and travel time unsatisfactory, but the evaluations increased steadily; while they perceived safety, comfort and convenience as satisfactory, but the assessments declined moderately over time.

Moreover, contributing to the longitudinal analysis of the public's acceptance of driverless shuttles, a structural equation model is employed to explore the changes in the judging criteria regarding service adoption among adopters and non-adopters. Our findings indicate that the evaluation of comfort and convenience are the most significant determinants of satisfaction and perception of usefulness, which in turn positively affect people's adoption intention as well as their favorable word-of-mouth behavior. In particular, the initial evaluation of the shuttle service is essential for consistent acceptance of the new public transportation mode.

Additionally, current public transport users and tech-savvy people are identified as the potential user groups who are more likely to adopt the shuttle service in the future. It is expected that the provision of faster, safer, more comfortable and convenient riding experiences with driverless shuttles will eventually activate users' behavior intention and their word-of-mouth behavior.

Keywords: driverless shuttle, user acceptance, behavior intention, service quality evaluation, three-wave panel analysis

Acknowledgement

I would like to thank my thesis supervisor Jia Guo at first. Thank you for patiently answering all my questions. Even on the holiday or weekend, I can receive your reply and feedback with an amazing speed, I really appreciate that. Under your guidance, I have not only gained a deeper understanding of systematic and academic research methods but also learned how to think in a more rigorous and logical way.

Second, I am so grateful for the help from Rafael. It was so kind of you to take time out of your personal time to give me advice and technical supports. Although we are far apart now, you have contributed more than half of my precious memories and spiritual support in my whole postgraduate study career. To say thank you is not enough.

Lastly, I would like to express my gratitude to Shufan Ye, Jiayan Qiu, Yan Cai and my beloved Muqing Zhu. These six months were filled with frustration and moments of despair. It was you who comforted me, encouraged me and stayed with me until I finished the last sentence of this thesis. It was a challenging time, but you made it easier.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DS: Driverless Shuttle

ADS: Automated Driving System

ODD: Operational Design Domain

DDT: Dynamic Driving Task

ADS-DV: Automated Driving System-Dedicated Vehicle

LSEM: Longitudinal Structural Equation Models

LGCM: Latent Curve Growth Models

1. Introduction

1.1 Research Background and Motivation

During the last decade, autonomous vehicles (AVs) related topics have been quickly developed from emerging to being frequently researched worldwide. Especially the automated driving technology has the potential benefits to improve the current transport system. According to macroscopic traffic flow theory, the wide pervasive use of automated vehicles can significantly improve road capacity, thus bringing about more efficient utilization of road infrastructure (Friedrich, 2016). The benefits of AVs also include improving road safety by mitigating traffic accidents due to human error (Petrović et al., 2020) and reducing pollutant emission along urban corridors in the context of mixed traffic by reaching a certain penetration rate of electric AVs (Rafael et al., 2020; Tomás et al., 2020).

Nevertheless, as discussions and research were going on, scholars realized that private automated vehicles might not solve road congestion in a couple of decades. In contrast, it will increase the total number of inflow vehicles on the road due to the low willingness to share a ride. In the worst case, it may cause traffic congestion to increase in the short term before easing in the long term (Alexander-Kearns et al., 2016; Kellett et al., 2019). Alternatively, as a public transport service, by improving the occupancy rates of vehicles and reducing the total amount of vehicles on the road, shared automated mobility could help to reduce traffic congestion and make future urban mobility more sustainable and efficient.

Over the past years, due to the restrictions of laws and regulations, fully autonomous vehicles were not permitted to operate in the mixed traffic context together with pedestrians, bicycles and social vehicles in the real on-road environment. Therefore, it was limited to collecting data on people's reactions in a simulated environment in the test fields. The ideal conditions provided by the specific experimental site (e.g. closed campus) inevitably led to the overestimated positive response biased from real-world conditions. In the case of a random sample of the population, the scholars could only distribute questionnaires to people without the actual physical experience, collecting expected values of indicators regarding the intention to use shared autonomous service under hypothetical scenarios. Using such hypothetical scenario-based approaches, it is rather difficult to examine people's actual opinions towards the automated shuttles as users are not exposed to the real bus service, leaving knowledge gaps for further studies. Among the few notable exceptions, several ongoing studies address the importance to collect attitudes and perceptions from respondents who have ridden automated shuttles in the real-life environment instead of the hypothetical environment (Nordhoff et al., 2019a; Guo et al., 2020; Chee et al., 2021).

With more riding experiences and knowledge of vehicle automation technology, individuals' attitudes towards the new bus system will change over time (Nordhoff et al., 2019a; Chee et al., 2021). Consequently, to successfully switch from the demonstration stage to the popularization of its regular use of driverless shuttle as a feeder service, it is imperative to continuously monitor the attitude factors and service quality attributes affecting user acceptance and their changing trends. Based on the data collected from pilot demonstrations and operation trials on the open and mixed public road, a considerable number of studies have explored the factors influencing user acceptance of driverless shuttle with real-life ridden experience. However, most of the existing research evidence is based on cross-sectional data and thus has no data to explore whether attitudes would change over time. To fill the research gap, a longitudinal survey was conducted. Using the panel data, the present study focuses on the real-world riding experiences of the fully automated public bus service operated in Stockholm. We expect the findings to provide information to better understand the public's attitudes change towards the new bus mode and whether attitudinal changes vary with socio-demographic characteristics.

1.2 Objectives and Research Questions

This study aims at filling the gaps in the longitudinal analysis of intention to use the driverless shuttle as a feeder to mass transit and further recommend the service to others, while understanding the changing trends of perceived service quality. According to attitudinal changes from the panel survey conducted within the same sample at three time points, the intra-individual and inter-individual variability pattern will be drawn. Hereby, the current study is part of the modern mobility in Barkaby (MMIB) project, as one of the research themes under Integrated Transport Research Lab (ITRL; KTH, 2020). It validates and refines the antecedent conceptual framework of initial and continuance intention influenced by perceived service qualities (Chee et al., 2021), which was based on the panel questionnaire survey investigated at another pilot project site, Kista, Stockholm. At the same time, this thesis enriches the case study of longitudinal user acceptance modeling based on the precursor research conducted by Guo et al. (2021) using an extended TRA model under the context of driverless shuttle service. Eventually, this study will provide a longitudinal view of how people perceive and tend to adopt driverless shuttle even recommend the service to others.

The following research questions are going to be explored to achieve the objectives of this thesis:

(1) How the intra-individual changes and inter-individual variability in evaluations of the driverless shuttle service develop over time?
(2) What does the temporal pattern of longitudinal user acceptance of the shuttle service look like? Will the determinants of behavioral intention change over time?
(3) How the different individual characteristics influence the perceptions, evaluations and adoption intention of the driverless shuttle service?

1.3 Expected Contributions

This thesis is expected to contribute to exploring the longitudinal changes in the judging criteria regarding the driverless shuttle service adoption among adopters and non-adopters. Concretely speaking, the thesis is expected to contribute to the existing literature to the following aspects:

(1) This study is expected to firstly provide a deeper understanding of the longitudinal changes in the public's attitudes and perceptions of the driverless shuttle service over time. The study will reveal the general changing tendency of people's perceptions of three service attributes (i.e. functional, safety and hedonic attributes) that contributing to the overall service satisfaction, facilitating the improvement of satisfaction in the

next operational phase. The results are also expected to explain the interindividual variability in the average tendency among respondent groups with different personal characteristics and varying socio-demographic profile, providing evidence for the development of targeted tactics.

(2) Subsequently, this study is expected to probe people's evaluation criteria in relation to the assessment of adopting the driverless shuttle service and recommending it to others. The time-varying determinants of adoption intention will help to promote the adoption rate by enhancing the specific aspect of the service in different stage. At the same time, it will be easier to create a favorable atmosphere of public opinion through a word-of-mouth marketing strategy.

(3) Ultimately, it is expected to demonstrate the temporal effect of prior assessment on continuance intention of adoption and recommendation, providing a reference for establishing a larger and more stable user group in the future.

1.4 Research Framework and Organization of the Thesis

The research framework is displayed in Figure 1. In line with the defined research objectives and research questions, this thesis is organized into six chapters. The first chapter of the thesis introduces the research motivation and objectives. The second chapter gives an overview of essential concepts and existing relevant studies. Subsequently, the methods used in this study are introduced, which includes latent curve analysis and longitudinal structural equation modeling. Chapter 5 presents the estimation results and comprehensive analysis. Last but not least, the conclusion is drawn in chapter 6, followed by the discussion, limitations and future directions of this study.

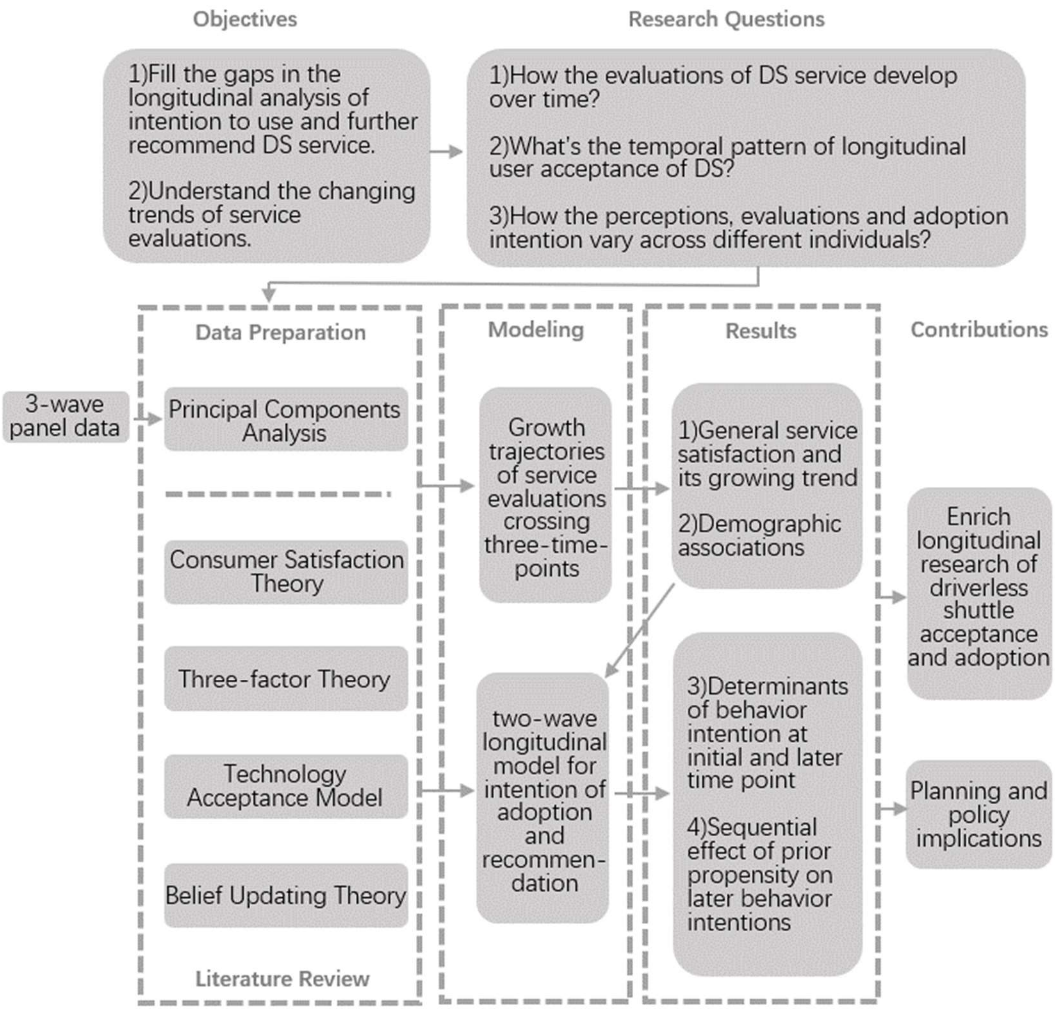


Figure 1 Research Framework

[this page is supposed to be blank]

2. Literature Review

This chapter aims to gain an understanding of the existing research regarding public acceptance and adoption of shared automated vehicles and introduces relevant concepts of word-of-mouth behavior, service evaluation and users' satisfactions. Section 2.1 distinguishes concepts of automated vehicles, and Section 2.2 provides a comprehensive review of the current driverless shuttles demonstrations and corresponding public's attitudes toward the shuttles. Next, Section 2.3 reviews the well-established user acceptance theories and corresponding empirical applications in the driverless shuttles context. Subsequently, Section 2.4 recalls the importance of word-of-mouth behavior for the acceptance of driverless shuttle service. Lastly, Section 2.5 highlights that quality of service and satisfaction with the service are also important factors contributing to driverless shuttles' adoption.

2.1 Automated Vehicles

With the growing application of vehicle automation technology in the transportation discipline, Automated Driving System (ADS) has been successfully applied in different transportation modes. A wide variety of automated mobility solutions have emerged, including private and shared automated vehicles (driverless shuttle bus, self-driving car-pooling and ride-hailing, driverless taxi and so on). Developing jointly with the mobility as a service (MaaS) concept, these emerging automated mobility solutions will definitely result in another earth-shaking revolution for our travel modes in the future.

Yet, since the ever-evolving industry is still in the initial launching stage, various vehicle manufacturers and technical support parties have appeared in the market. Besides, both the hardware and software of the automated vehicles have been upgraded rapidly. Therefore, discordant appellation and naming methods have been used by various research institutes in different fields. Notably, it is prevalent that the terminologies “self-driving”, “driverless”, “automated” and “autonomous” are interchangeably used without differentiation. However, there are unignorable hidden differences between the terminologies. First, driverless vehicles do not have a human driver on board as the fallback to take over the control when self-driving vehicles encounter unforeseen incidents (David Levinson, 2017). Second, autonomous vehicles commonly refer to the automation level above Society Automotive Engineers' (SAE's) standard (On-Road Automated Driving (ORAD) committee) level 4, but an automated vehicle might need human monitoring or intervention (Adler, 2019). Driverless and autonomous were deemed nearer to synonym, which was concluded by Levinson (2017).

For a more rigorous and detailed definition and classification of the existing edge-cutting automated driving products, SAE's latest revised taxonomy standard J3016 can be referenced. It is worth mentioning the given definition of Automated Driving System-Dedicated Vehicle (ADS-DV): "An ADS-DV is a truly "driverless" vehicle. It might be designed without user interfaces; also might be operated temporarily by a conventional or remote driver." The research object of this paper is precisely in line with the above definition, applying SAE's Level 4 ADS without a user interface.

In addition to the degree of automation level, the variety of autonomous mobilities can be grouped by different passenger capacity. (Ainsalu et al., 2018) summarized features of the existing automated mobility solutions compared to conventional transport modes with multi-standard classification. From the perspective of occupancy of the vehicle, modes of transport can be categorized into riding alone (occupancy=1 passenger), riding with guest (occupancy=2-5 passengers), group transit (occupancy=5-10 passengers) and mass transit (occupancy=10-200 passengers). Corresponding to each occupancy range, the implementation of ADS in different modes can be defined as "self-owned autonomous automobile", "self-driving car-pooling and car-hailing", and "driverless transit". Although the application of ADS to mass transit are being filled with various ongoing projects, the automation level of them still remains at SAE's level 3, such as the full-length automated bus project in Sweden (TRATON GROUP, 2019), the autonomous electric bus trial jointly launched by Volvo and Nanyang Technological University (Volvo Buses Global, 2019), the 5G self-driving bus trial in Zhengzhou, China (YUTONG, 2020), and Europe's first trial of a full-sized, autonomous passenger carrying bus on a regular basis in Málaga, Spain (SPACE, 2021). The first two solutions are less efficient, and currently, there is no fully autonomous solution to mass transit. Consequently, driverless shuttles are considered the most efficient driverless mobility solution with the largest vehicle occupancy.

Taken together, according to the automation level and occupancy of the vehicle, the term of the current research object can be clarified. In order to prevent either non-uniform or esoteric naming from confusing non-peer and general audiences, and to distinguish from full-length autonomous mass transit equipped with standard manual controls, the term "driverless shuttle (DS)" is adopted by this thesis. This term is in line with the definition of ADS-DV given by SAE and has been used in several studies discussing public acceptance and intention to use the service. (Smolnicki and Sołtys, 2016; Sina Nordhoff et al., 2017; Ainsalu et al., 2018; Salonen, 2018; Mahmoodi Nesheli et al., 2021)

2.2 Driverless Shuttles

An abundance of works pertaining to various aspects of the driverless shuttle is emerging worldwide. The vast majority of studies were based on completed or ongoing demonstrations in European countries, including Germany, France, Switzerland, the Netherlands, Finland and Sweden. Such a vigorous development is because EU funded project CityMobile2 has primarily removed the barriers to launch the driverless shuttle service on public roads in Europe, where driverless shuttles were positioned as an integral part of the automated road transport system (ARTS). The project results reveal that the most decisive factor affecting the public's attitudes towards fully automated mobility is perceived safety. ARTS's safety requirements also turned into integrating the in-vehicle security perception and interaction between the ARTS vehicles of ARTS' users and other road users. The operating speeds were limited to around 10km/h, allowing for driving in segregated lane or non-motorized areas shared with pedestrians and bicycles to ensure safety. Such limited condition with high probability conforms to the infrastructure and traffic configuration of supplementary branches in local road networks or traffic calm areas in residential areas. This inference is consistent with the conclusion in the CityMobil2 final report (2014), which points out that the most supported role of the autonomous bus was considered a transport framework supplement.

The positioning of the driverless shuttle also reflects the travel demand pattern of modern people. There has been a trend to combine multiple transportation modes to complete a single trip, with the last mile to and from the transit hub being the most publicly accepted part of the seamless travel chain. In the Norwegian project Smartfeeder (2017), one-fourth of the respondents thought there was a need for a driverless shuttle when the shuttles running between residential and public transit (Roche-Cerasi, 2019). Experienced people also gave a positive response to the idea of using the driverless shuttle as feeders to public transit in a semi-structured interview in Germany (Nordhoff, Winter, et al., 2019). It is convinced to conclude that by improving the accessibility of public transport station and extending mobility to all, driverless shuttles would be easier accepted by the public as a feeder service. As the last mile solution, trying to replace driving private cars and car-hailing; or becoming an alternative to shared bikes and e-scooter sharing is the most suitable key breakthrough for the driverless bus. Therefore, the role of the driverless shuttle in this study has been supported and validated, which is in line with the holistic EU strategy.

So far, as summarized in the final report of the CityMobil2 project (2014), people show relatively positive attitudes in the demonstration cities, Oristano (Italy, July and August 2014), La Rochelle (France, from October 2014 to April 2015), Lausanne, (Switzerland, from October 2014 to April 2015), Vantaa (Finland, in July and August 2015), Trikala (Greece, from August 2015 to February 2016), Sophia Antipolis (France, from March to May 2016) and San Sebastian (Spain, from April to July 2016), where driverless shuttles served as feeders to public transport. Likewise, Bansal et al. (2016) found 41% of the respondents in Austin prefer to use shared automated vehicles weekly at a competitive price. However, it has been demonstrated by Becker and Axhausen (2017) through a review study that the general public opinion is not unanimously positive but can vary with many factors, such as perceived usefulness, safety, comfort, reliability and so on. Site et al. (2011) suggested that a high on-board comfort was expected by passengers who tend to use the fully automated shared vehicle in a stated preference survey. In particular, the elderly expected a higher riding comfort as benefits from automation. Moreover, respondents in the interview study addressed by Nordhoff, Winter, et al. (2019) deemed convenience as a driver encouraging them to adopt driverless shuttles in public transport. Nevertheless, a case study in Vanta (Finland) reported 64% of driverless shuttle passengers perceived in-vehicle security of the driverless shuttle worse than in the regular bus (Salonen, 2018). Additionally, reasonable travel time and waiting time were proposed to be the critical determinants of the use of shared automated vehicles (Krueger et al., 2016). Despite these service attributes of driverless shuttles, demographic characteristics diversify people's perceptions of automated vehicles. These demographic associations were reviewed by Nordhoff, Kyriakidis, et al. (2019), implying that men, younger people, those living in dense urban areas and people who are more familiar with innovative technology are more accepting of automated driving products.

Bring it further, from an overview of studies to date (Azad et al., 2019), themes related to driverless shuttle can be roughly divided into five categories： 1) technology capability; 2) user acceptance; 3) security and safety concerns; 4) social and economic impacts; 5) policies and regulatory issues. Since only by being accepted by most travelers and maintaining positive feedback on the service, can a driverless shuttle be successfully and widely deployed (Alessandrini et al., 2014; Bansal et al., 2016; Nordhoff et al., 2019a; Guo et al., 2020; Chee et al., 2021). This thesis falls to focus on the customers' satisfaction and their evaluation criteria over time.

2.3 Driverless Shuttles and User Acceptance

Explaining user acceptance of new technology is one of the most mature fields of study in the information technology industry. Within the diverse variety of theories, one stream of research uses intension or usage as a dependent variable to analyze individual acceptance. Up to now, eight mainstream models have been established, developed and validated. Although the model structure and supporting theories are different, they share a common basic conceptual framework (Figure 2). As shown in the figure, people's intention of use is influenced by their reactions to using the system, which would subsequently influence their actual behavior (Venkatesh et al., 2003). The unified theory of acceptance and use of technology (UTAUT) sheds light on the conceptual underpinnings of these models that they could be unified. Moreover, using a dataset over a six-month period with three points of measurements, Venkatesh et al. (2003) discussed and compared eight models. The eight models reviewed are integrated theory of reasoned action (TRA), technology acceptance model (TAM), motivational model (MM), theory of planned behavior (TPB), combined TAM and TPB (C-TAM-TPB), model of PC utilization (MPCU), innovation diffusion theory (IDT), social cognitive theory (SCT). After adapting the unique predict factors in the eight models to the same dataset, the explanatory power of each variable to behavioral intention was sorted. Eventually, a generalized model is obtained after unification and simplification, in which the retained independent variables are performance expectancy, effort expectancy, social influence, facilitating conditions. This model integrates the core elements of eight competing models and provides a generalized way to assess the likelihood of success when introducing innovative technology to the public.

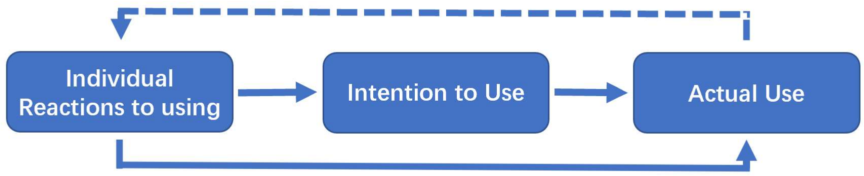


Figure 2 Basic Conceptual Framework of User Acceptance Models

As an innovative technology, understanding user acceptance of AVs can help to avoid unfavorable perception in the initial implementation stage, which could bring long-term negative effects onto the future adoption intention. For example, Osswald et al. (2012) proposed the car technology acceptance model (CTAM), which pioneered the application of the UTAUT model in the automotive context. The authors pointed out that the common deficiency during adaption is the negligence of adjusting explanatory variables according to the contextual situation. The hurdles in converting from the original IT environment into a car-driving context were also implied by the decreasing explained variance of the behavioral intention of use. For instance, UTAUT model (Venkatesh et al., 2003) achieved an explanatory power of 70% in the original IT environment but decreased to 20% in the car-driving context. The reason is argued by CTAM that some factors will become less significant due to changing the model environment and need to be discarded; Other unique variables need to be added to improve the interpretation of the model. In another behavior model, the first empirical application of the user acceptance model in the driverless shuttle context was under the Citymobil project. Using the UTAUT model, (Madigan et al., 2016) compared ARTS users' adoption intention in La Rochelle, France and Lausanne, Switzerland. Their study proved that the model could be applied to group transit, and laid the foundation for further exploration of further factors influencing acceptance of the driverless shuttle. However, similar to the situation when applying CTAM, the explained variance in behavior intention in this case study also reduced to 20% when the contextual situation switched from IT to the driverless shuttle. Madigan also argued that more emphasis should be placed on system performance rather than looking for the best way to use the system with increasing user performance of a task in the driverless shuttle context. Therefore, it is essential to tailor the measurement construct and its corresponding indicators to reflect the context in which respondents interact with the driverless bus. In addition, performance expectancy is the most powerful explanatory variable among all the factors, which has also been verified by the later case studies in Berlin-Schöneberg and the recent EMMA project in Mainz.

Table 1 literature review of previous studies regarding user acceptance of driverless shuttle

Route	Route length	Pilot site	Method	Factor	Moderator	Sources
A popular tourist route	1.7km	La Rochelle (France)	UTAUT	Perceived Expectancy(PE) Effort Expectancy(EE) Social Influence(SI)	age, travel behavior	(Madigan et al. 2016)
A link between the metro station and main working sites	1.6km	Lausanne (Switzerland)	Adapted UTAUT	Perceived Expectancy(PE) Effort Expectancy(EE) Social Influence(SI)	age, travel behavior	(Madigan et al. 2016)
A link from railway station to the exhibition centre including 100m tunnel	0.9 km	Vantaa (Finland)		Safety Security Emergency Management		(Salonen 2018)
A loop route in city centre cross heavy traffic road, city street and two bridges	2.4km	Trikala (Greece)	Adapted UTAUT	Perceived Expectancy(PE) Effort Expectancy(EE) Hedonic Motivation(HM) Facilitating Conditions(FC)	age gender experience	(Madigan et al. 2017)
A route on EUREF office campus	0.7km	Berlin-Schöneberg (Germany)	UTAUT	Shuttle and Service Characteristics Shuttle Effectiveness	gender, age	(Nordhoff et al. 2018)
A straight line in the pedestrian area	0.6km	Mainz (Germany)	Adapted UTAUT	Perceived Expectancy(PE) Effort Expectancy(EE)	age, gender, experience	(Bernhard et al. 2020)
A last-mile operation line to the train station	0.8km	Stockholm -Kista (Sweden)		Service Attributes: Safety Ride Comfort Travel Time Reliability		(Chee et al. 2021)
A regularly scheduled feeder line connects the residential area to the transit hub and shopping centre	2.5km	Stockholm - Barkarby (Sweden)	Extended TRA	Attitudes towards Behavior(ATB) Social Norms(SN) Needs		(Guo et al. 2021)

Most recently, a comprehensive conceptual model, integrating empirical works to date in the ADS-DVs context, was proposed by (Nordhoff et al., 2016). Incorporating with UTAUT3 and TAM, subsequently, the conceptual framework was refined by (Nordhoff, Kyriakidis, et al., 2019) as a more mature multi-level model on automated vehicle acceptance (MAVA). The conceptual framework, in particular, focused on the driverless shuttle functioning as a feeder to the public transit system. Acceptance was treated as a 4-stage decision-making process, starting with exposure to ADS-DVs and followed by attitudes towards ADS-DVs ending at the intention to use and then actual adoption of ADS-DVs. Specifically, stage 2 was subdivided into three aspects: 1) instrumental domain-specific; 2) symbolic-affective; 3) moral-normative characteristics of ADS-DVs. Additionally, socio-demographics, travel behavior and personality were included in the model at the micro-level to explain the individual difference.

A growing body of studies explored influential factors in user acceptance of AVs (see Table 1). However, most of the previous studies were based on cross-sectional data, lacking understanding of the temporal changes in the public's opinions toward driverless shuttles, leaving knowledge gaps for further studies. Among the few notable exceptions, (Chee et al., 2021) drew attention to the longitudinal changes in users' valuation of a driverless shuttle trial in Stockholm-Kista. In the study, Chee et al. postulated a structural equation model to explain adopters' changing concerns affecting their continued adoption intention with increasing ride experience. The results implied that adopters' intention to adopt the shuttles were initially influenced by safety and travel time, whereas ride comfort became a primary determinant with increasing ride experiences. However, these findings were based on a demonstration trial where a driverless shuttle service only operated along an 800 meters straight line, lacking validation in a mixed traffic environment on public roads. Therefore, based on a panel survey conducted in Stockholm-Barkarbystaden, the current study aims to fill this research gap and enrich the longitudinal analysis of the initial and continued adoption intention of a regular-scheduled driverless shuttle service.

2.4 Word-of-mouth Recommendations

In the precursor work, word-of-mouth (WOM) recommendation was involved in the user acceptance model proposed by Guo et al. (2021; listed in Table 1). Applying an extended theory of reasoned action model, their findings conveyed that attitudes, social norms and perceived usefulness positively impact not only behavior intention but also WOM behaviors. This is in line with the determinants of WOM in the marketing field, which contains consumer satisfaction, social tie strength, perceived novelty and so on (Bone, 1992). The concept of word-of-mouth communication was derived from the marketing industry. It refers to verbal communication about products and services, from consumer to consumer (W.H. Whyte, 1954; Robert M. Schindler and Barbara Bickart, 2005). The salient role of WOM has been widely acknowledged in forming consumers' attitudes and behaviors (Brown and Reingen, 1987). WOM recommendations are given by satisfied clients or potential adopters with positive attitudes and spread to their acquaintances (Harrison-Walker, 2001). These influencers would subsequently create an exposure environment for new potential adopters. Positive WOM behavior functioning as an external stimulus would help speed up the adoption rate (Harrison-Walker, 2001; Nordhoff et al., 2019). This reinforcing loop of WOM in AVs context also was highlighted from the system dynamics perspective (Nieuwenhuijsen et al., 2018; Berg et al., 2020). By considering WOM during the commercialization process of the driverless shuttle, a broader view of public acceptance can be obtained.

2.5 Quality of Service and Satisfaction

On the contrary to AV, shared automated mobility performs more in the manner of public transport. As the result of a business to customer (B2C) commerce mode, the pattern of users' perception and intention to use driverless shuttle differ from the ones of self-owned AV. For public transit passengers, the satisfaction or expectation of service quality plays a paramount role when they make the decision on mode choice (Stocker and Shaheen, 2017). Before the current study, a pilot project carried out in Stockholm-Kista, which was also administered under the Integrated Transport Research Lab (ITRL). Chee et al., (2020) applied three-factor theory analysis to identify user's core perceptions affecting their using intention of the driverless shuttle service. The study substantiated the existence of Maslow's hierarchy of transit needs in the acceptance of driverless shuttle context. The framework of Maslow's hierarchy is shown in Figure 3, where transit needs have been adapted to understand public transport satisfaction (Allen et al., 2019). Functional attribute, security attribute and hedonic attribute were concluded as three aspects compromising the triangular hierarchy, which are in line with three-factors theory (KANO et al., 1984; Allen et al., 2019; Chee et al., 2020). More specifically, three factors refer to the basic, performance and exciting factor of service quality (KANO et al., 1984; Zhang et al., 2015; Wu et al., 2018; Abenoza et al., 2019). Corresponding characteristics (e.g. travel time, safety, comfort, etc.) of the driverless shuttle were classified and validated by (Chee et al., 2020).

When introducing driverless shuttles to the existing transport system, it is necessary to evaluate the impact of service quality on public acceptance and usage of the new transportation mode. Nordhoff et al. (2018) conducted a study that interviewed 30 users after the actual ride of a driverless shuttle on the EUREF office campus in Berlin-Schöneberg to understand users' acceptance of the shuttles in an in-depth way. Their study highlighted that quality of service is one of the essential determinants of driverless shuttle acceptance. Service quality was widely accepted as one of the antecedents of satisfaction (Fu and Juan, 2017; Allen et al., 2019; Oña, 2021). Furthermore, the mediator effect of satisfaction between service quality and behavior intention suggested that service quality attributes might indirectly influence intention to use the driverless shuttle (Fu & Juan, 2017; Oña, 2021). Last but not least, the determinant role of service quality on use intention was also validated in the MAVA (Nordhoff et al., 2019).

Thus, to enrich the understanding of service quality's mediating role in the user acceptance of driverless shuttles, this study aims to verify the decisive effects of service quality evaluations on overall passenger satisfaction, which in turn affects adoption intention.

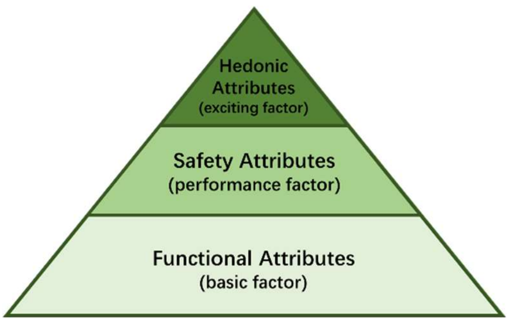


Figure 3 Triangular Service Quality Attributes Hierarchy

3. Data Collection and Analysis

3.1 Study Area

This thesis is part of the modern mobility in Barkaby (MMiB) serial studies, under the Integrated Transport Research Lab (ITRL) at KTH Royal Institute of Technology. The research target of MMiB is the driverless shuttle *Line 549* in Stockholm-Barkaby, Sweden, which is jointly developed by local transport administration Stockholm Public Transport (SL), Järfälla municipality and the operator Nobina. The project aims at supplying low car-dependent, seamless and intelligent mobility for all user groups in the most prominent northern European new-developing residential area.

Barkabystaden is a sustainable and smart city project parcel with expected 40,000 inhabitants in western Stockholm (Barkarby, 2020). The Barkarby model has been built, which collaborates urban development with infrastructure, public transport and innovative technology to inaugurate an attractive and modern neighborhood. By 2026, a planned electrified BRT route and on-demand driverless shuttle feeder will provide unbeatable accessibility to make Barkarby an important future transport hub.

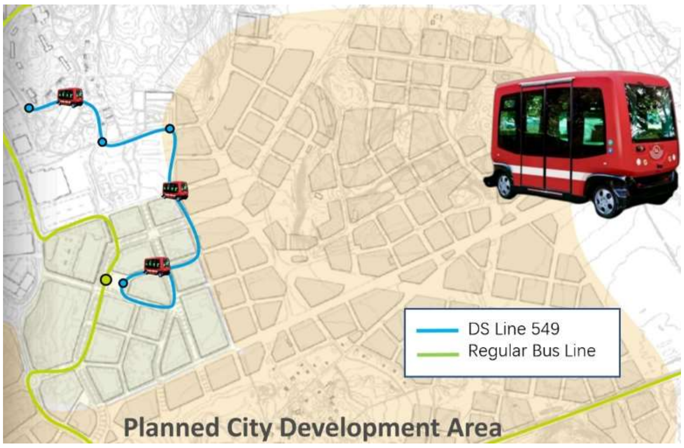


Figure 4 Barkarbystaden Driverless Shuttle Line 549 Operation Route.

Source: https://www.drivesweden.net/

In this direction, the world first regular scheduled driverless shuttle *Line 549* hits the road in October 2018. Three electrically powered EZ10 vehicles (EasyMile) served *Line 549.* They could carry up to 11 passengers per vehicle at a time, operating during peak hours between 06: 30 and 09: 00, and between 12: 30 and 19: 00 on weekdays, and between 11: 30 and 19: 00 on Saturdays. The length of the route is 2.5 km, which connected two newly built residential areas, Herrestaskolan and Stora Torget (see Figure 4). The maximum speed is 15 km/h, and it took 11 minutes to finish the whole route. The service has four stops and could be accessed by SL tickets. After the first phase of the trial, it will be tested for on-demand mode with doubled route length in the forthcoming phases. During the trial, an onboard steward was backing up for the emergency by holding a portable control panel.

3.2 Data Collection

Three waves of panel survey were conducted in March, September, and December 2019. An online survey company recruited inhabitants and workers in the research area as respondents who might have opportunities to interact or use the driverless shuttle.

Questions used in this thesis could be grouped into four sections. The first section asked question regarding the degree of interaction between the respondents and the driverless shuttles. Respondents were asked whether they have seen the driverless shuttles by themselves first. Then they were required to report the cumulative times of using the service and times of riding in the last one week before the interview. The second section answered a series of five-point Likert-scale questions related to public attitudes and perceptions of driverless shuttles, including the functional service characteristics (e.g. travel time, speed, frequency etc.), evaluations and expectations of user experience performance (e.g. convenience and onboard comfort) and perceived safety. The third section asked questions regarding respondents' use intention, willingness to recommend, satisfaction and perceived usefulness of the service. The final section assessed respondents' demographics, along with travel behavior and personality. In this study, travel behavior refers to accessibility to private mobility (i.e. car ownership) and weekly use frequency of transport mode as a first-/last-mile solution connecting the nearby transit hub. Likewise, Personality was presented as technology savviness in the technology concern section of the questionnaire.

3.3 Data Description

After approximately nine months of the three-wave panel survey period, the number of completed responses were 511 in the first wave, 445 in the second wave and 584 in the third wave, respectively. Among all those responses, only 363 cases with fully complete three wave answers are reported here. Besides, the response of behavior intention and word-of-mouth recommendation intention were only collected in the last two waves.

Socio-economics characteristics

Table 2 Scio-economics characteristics of the respondents

Variable	Coding	Classification	All Respondents (N=363,%)
Gender	1	Male	50%
	0	Female	50%
Age	1	Youth (15-24 years old)	4%
	1	Adult (25-44 years old)	59%
	0	Middle-age (45-64 years old)	24%
	0	Elder (⩾65 years old)	12%
Education	1	Higher than/qual to Bachelor	45%
	0	Lower than Bachelor	55%
Employment status	1	Employed	79%
	0	Unemployed	1%
	0	Student	7%
	0	pensioner	12%
Annual income level	1	Low-income (<30,000€)	12%
	2	Middle-income (30,000-69,000€)	42%
	3	High-income (⩾ 30,000€)	25%
	-	Do not want to specify	21%
Car ownership	1	Yes	76%
	0	No	24%
Live near the DS line	1	Yes	94%
	0	No	6%
Tech-savvy	1	Yes	89%
	0	No	11%
Familiarity with autonomous driving technology	1	Yes	21%
	0	No	79%

In total, 363 respondents completed all three waves of the survey. Among them, 131 people had taken driverless shuttle at least once, and 232 people had no experience with the shuttle service at all. As presented in Table 2, among all the respondents, males and females were equally distributed. The uppermost age span among the sample was the adult group (25-44 years old) with a share of 59% and the middle-aged group (45-64 years old) with 24% of the sample. With regard to education level, high-educated people were slightly dominant, which means that respondents with a Bachelor degree or higher than a Bachelor degree occupied 45% of the sample. In terms of employment status, 80% of respondents are employed, followed by pensioners with 12%. Regarding income level, people with upper and middle incomes were the largest income classes. It is worth noting that 21% of respondents were not willing to disclose their income level. Additionally, 76% of respondents owned at least one car in the household. Lastly, since the questionnaire was distributed in Barkarby, Stockholm, 94% of the respondents stated they live along the driverless shuttle service line.

Moreover, individual preference for innovative technology was investigated by asking about technology savviness and familiarity with the technology behind the driverless shuttle. As shown in Table 2, although 89% of respondents self-rated themselves as tech-savvy people, only a quarter are aware of the technology supporting the driverless shuttle to run on the road without a human driver. Thus, the public's knowledge and awareness about autonomous driving technology still can be improved. To understand the technology savviness item from a more specific viewpoint, respondents' attitudes toward technology was measured by five well-designed questions in the third wave questionnaire. As shown in Table 3, all five items were measured in the seven-point Likert scale, ranging from completely disagree to completely agree. The results reveal that 79.1% of respondents have relatively high technological openness. At the same time, 63.7% of respondents agreed on the importance of following the technology development. The last two items at the bottom of Table 3 measured the enthusiasm for technology. Results indicate that a fair number of respondents (approximately 40%) self-assessed that they have more curiosity and desire for innovative technology than others. Lastly, around 30.8% of the sample had confidence in their knowledge of the most cutting-edge technology products. In general, acceptably portion of people willing to embrace the change and have a positive attitude towards technology.

Table 3 Attitude towards Technology Descriptive Statistics

Statement	Question ID	Response (N=363,%)						
		Completely disagree	Disagree	Partially disagree	Neutral	Partially agree	Agree	Completely agree
I think that new technology creates more problems than it solves.	W3Q12.1	33.1%	28.4%	17.6%	9.6%	5.8%	3.0%	1.9%
It's important for me to follow the technology development.	W3Q12.2	5.8%	6.1%	11.3%	13.2%	21.8%	23.4%	18.5%
I'm always among the first to try to use new technology products.	W3Q12.3	15.2%	12.7%	17.9%	24.5%	16.3%	8.0%	5.2%
I'm excited about the possibilities with new technology.	W3Q12.4	14.9%	11.8%	14.6%	17.6%	21.5%	8.5%	10.2%
I know more than others about the latest technology products.	W3Q12.5	22.0%	13.8%	11.3%	21.5%	14.6%	12.1%	4.1%

Awareness and usage of the automated shuttles

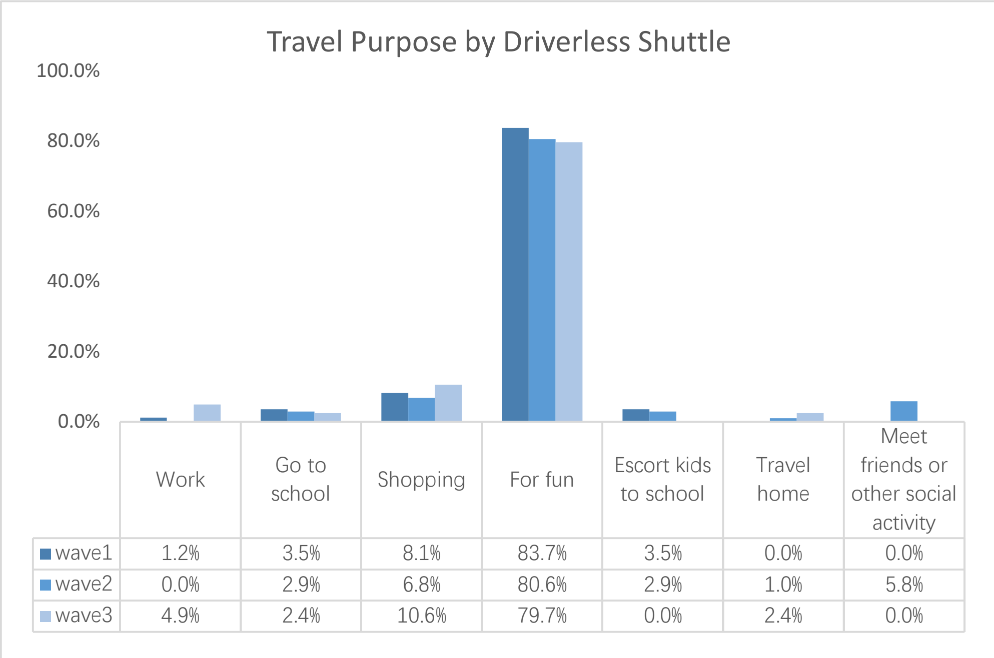
As shown in Table 4, awareness and usage of the driverless shuttle were reported. The results show that 87.6% of respondents had seen driverless shuttles in Barkarby in the first wave, and it was increased by 4.1%, reaching 91.7% in the third wave ($t=3.32$, $p<0.01$). In addition, we examined the adoption of DS. We can see that the portion of respondents with ride experience rose from the starting share of 24.1% to 36.1% at the end by a modest 2.1% ($t=13.04$, $p<0.01$). Although 131 respondents had taken the driverless shuttle by the end of the survey, 110 of them had no more than 5 times of ridden experience. Only 13 respondents ended up taking the service 6 to10 times, and 8 respondents used the service more than 10 times. The findings imply that it is urgent to enhance the continued adoption of the shuttle service. This study also collects information refers to how active the adopter was for the driverless shuttle service. Respondents reported the use frequency of driverless shuttle last week. Looking at the trends of last week use frequency over three time points, we see that the adopters remained in a comparatively low active usage status throughout the survey period. Despite the growing number of people adopting the service, still few converted to regular users after the initial attempt. In such a way, the majority of respondents adopted the driverless shuttle service only once. In order to veritably function as regular public transport, driverless shuttle service needs to retain first-time users and integrate this service into their daily travel patterns.

Table 4 Awareness and Adoption of Driverless Shuttle

Statement	Coding	Classification	Wave 1	Wave 2	Wave 3
			Count (%)	Count (%)	Count (%)
Have you seen DS in Barkarby?	0	No	45 (12.4)	35 (9.6)	30 (8.3)
	1	Yes	318 (87.6)	328 (90.4)	333 (91.7)
Cumulative times ridden DS	0	Never ridden	276 (76.0)	255 (70.2)	232 (63.9)
	1	1-5 times	75 (20.7)	91 (25.1)	110 (30.3)
	2	6-10 times	9 (2.5)	13 (3.6)	13 (3.6)
	3	10-15 times	2 (0.6)	2 (0.6)	5 (1.4)
	4	More than 15 times	1 (0.3)	2 (0.6)	3 (0.8)
Number of times ridden DS in last week	0	Did not use DS	348 (95.9)	353 (97.2)	346 (95.3)
	1	1 day	12 (3.3)	7 (1.9)	10 (2.8)
	2	2 days	2 (0.6)	2 (0.6)	4 (1.1)
	3	3-4 days	1 (0.3)	1 (0.3)	3 (0.8)
	4	More than 4 days	0 (0.0)	0 (0.0)	0 (0.0)

With respect to the travel purposes of the driverless shuttle trips among adopters, as shown in Figure 5, the study shows that by the end of the survey period, although riding for fun still was the uppermost motivation, it decreased from the initial 83.7% to 79.7% over time. 10.6%, and 4.9%, respondents reported their travel purpose as riding for shopping and commuting, ranking at second and third place among all the travel purposes, respectively. Therefore, it is possible to argue that novelty-seeking rides were slowly transforming into daily trips.

It is assumed that there was no significant change in the road network and traffic configuration in the study area during the nine months of the panel survey. As a consequence, it could be inferred that this trial operation has indeed successfully improved the public's cognition and the actual driverless shuttle rode experience in the research area. Additionally, the travel purpose for taking a driverless shuttle ride has been enriched and expanded as well.



	Work	Go to school	Shopping	For fun	Escort kids to school	Travel home	Meet friends or other social activity
wave1	1.2%	3.5%	8.1%	83.7%	3.5%	0.0%	0.0%
wave2	0.0%	2.9%	6.8%	80.6%	2.9%	1.0%	5.8%
wave3	4.9%	2.4%	10.6%	79.7%	0.0%	2.4%	0.0%

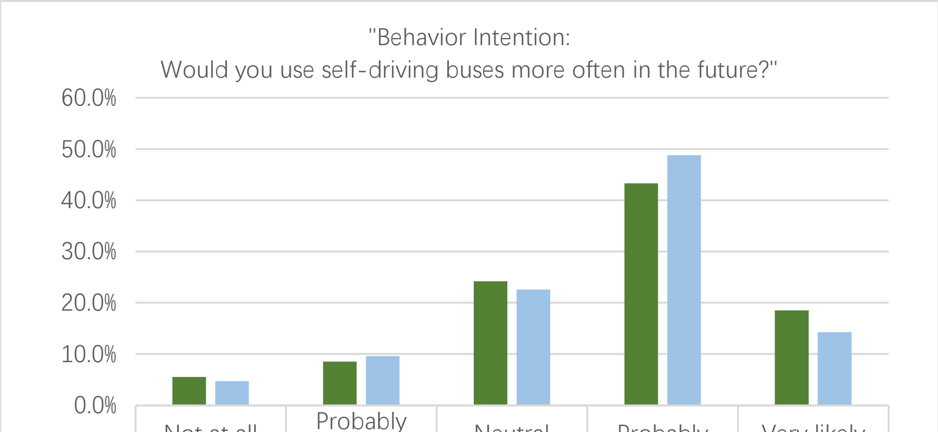
Figure 5 Travel Purpose by Driverless Shuttle

Behavior intentions and the word-of-mouth recommendations

The Public's behavior intention and word-of-mouth recommendation behavior were interviewed in the second and third wave surveys. These two variables were measured based on a Five-point Likert scale ranging from "Not at all" to "Very likely" by the following questions: "BI: Would you use self-driving buses more often in the future?" and "WOM: Will you recommend DS to your colleagues/friends/family?". As shown in Figure 6 and Figure 7, behavior intentions and word-of-mouth behaviors skew to the positive side in both stages. Moreover, only 5 to 6 per cent of respondents had strikingly negative responses to either BI or WOM. On the whole, the two variables were in line with the expectations.

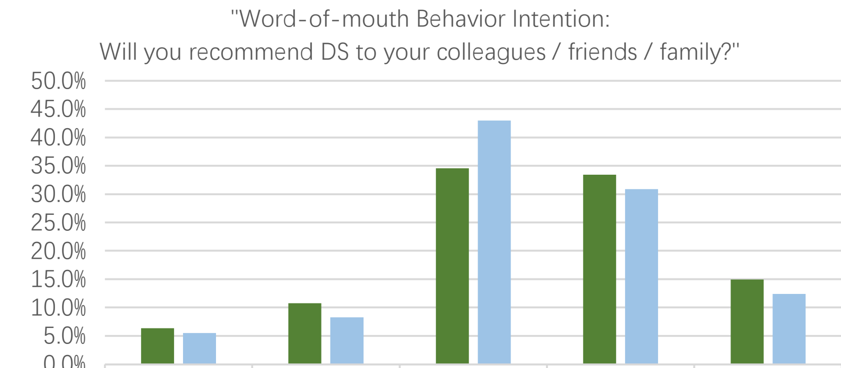
As the vital index of user acceptance, the results show that 62% of respondents would like to use DS services more frequently in the future. More specifically, the number of highly willing to enhance the future use of DS service and intensely resistant to keep using DS both declined, with 5% and 14% respectively, while willingness for future use "probably not" and "probably yes" rose to 9.6% and 48.8%, respectively. Neutral opinion had a barely observable slight reduction in the initial ratio of 24%. The mean scores of BI in the five-point Likert scale for the second and third wave were 3.61 and 3.58, but the differences were not statistically significant (t=0.464, p=0.643), which means changes between each scoring interval are mutually counteracted when the sample mean was taken. On balance, it is implied that respondents showed favorable attitudes toward DS service in general.

In terms of word-of-mouth recommendation, respondents were noticeably more conservative compared with BI. 35% of respondents were neutral in their recommendation willingness in the second wave, rising to 43% afterwards. It indicates that by the end of the survey, as many as 43% of respondents neither had a strong desire to recommend DS service to their relatives or friends nor strikingly expressed the propensity to refuse. On the contrary to BI, the number of participants who gave explicit attitudes declined, either positively or negatively. Ultimately, favorable attitudes at modest and extreme degree dropped to 31% and 12%; while the negative response at modest and extreme degree dropped to 5% and 8%. Similarly, the mean values of WOM in the two waves were 3.4 and 3.36, which did not differ significantly. Consequently, it elucidates that comparing mean values of BI alone cannot help us understand the mechanism and underlying factors contributing to the changes in individual attitudes towards the driverless shuttle. Different approaches need to be found to understand the nature of longitudinal user acceptance.



	Not at all	Probably not	Neutral	Probably	Very likely
BI2	5.5%	8.5%	24.2%	43.3%	18.5%
BI3	4.7%	9.6%	22.6%	48.8%	14.3%

Figure 6 Behavior Intention in Wave II and Wave III



	Not at all	Probably not	Neutral	Probably	Very likely
WOM2	6.4%	10.8%	34.5%	33.4%	14.9%
WOM3	5.5%	8.3%	43.0%	30.9%	12.4%

Figure 7 Word-of-mouth Recommendation in Wave II and Wave III

3.4 Data Preparation

As abundant attitudinal determinants were selected in this study, to reduce the dimensions of the data space and reconstruct representative items, a principal component analysis (PCA) was conducted. It is possible that the observed relationships of collected perceptions and evaluations might deviate from the assumed aggregates, leading to multicollinearity or low loadings on the corresponding component. Factor analysis can address these issues by examining the interrelationships between items, and its results could also provide evidence for the adjustment of the theoretical structure.

Prior to the analysis process, Kaiser-Meyer-Olkin (KMO) and Bartlett's test was conducted to measure how suited the data was for factor analysis. KMO values for the two waves were 0.78 and 0.70, respectively, which reach the middling sampling adequacy threshold of 0.7 (Stephanie Glen, 2016). Accordingly, the collected data was meaningful to conduct factor analysis for each independent variable and for the complete model.

Subsequently, eleven attitudinal variables were entered as input into the factor analysis process. After being extracted by the principal component analysis (PCA) method and rotated by the Varimax method with Kaiser normalization, three new primary components were obtained. The rotated matrix is shown in Table 5, where variables with communities less than 0.4 or eigenvalues smaller than 1.0 were omitted. Factor loadings in Table 5 reveal that, in general, variables in the second and third wave have consistent factor loadings and correspondingly belong to the same components. Furthermore, the proven consistency is a feasible basis supporting the follow-up longitudinal study.

Table 5 Rotated Component Matrix for Principal Component Analysis

Components	Statement of variables	Question ID	Component in Wave II				Component in Wave III			
			1	2	3	4	1	2	3	4
PC1: satisfaction or perceived usefulness of the shuttle service	Satisfaction of the driverless shuttle ridden trip	W2Q5 W3Q23	0.59				0.57			
	Prospect of the importance of the future role of the driverless shuttle in the transport system	W2Q17 W3Q28	0.80				0.89			
	Driverless shuttle would be suitable for daily travel.	W2Q22 W3Q31	0.83				0.81			
	Driverless shuttle would meet respondents' travel needs.	W2Q23 W3Q32	0.84				0.79			
PC2: perceptions of comfort and convenience	Speed of the driverless shuttle compared to a regular bus	W1Q23 W2Q27 W3Q36			0.74				0.81	
	Travel time of driverless shuttle compared to a regular bus	W1Q33 W2Q36 W3Q44			0.85				0.81	
	Travel time of driverless shuttle compared to a car	W1Q34 W2Q37 W3Q45			0.91				0.79	
PC3: perceptions of travel time and speed	General comfort of driverless shuttle trip	W1Q31 W2Q30 W3Q38		0.75				0.83		
	Convenience on board of driverless shuttle	W1Q24 W2Q31 W3Q39		0.84				0.83		
	Comfort on board due to driving speed and driving pattern of the driverless shuttle	W2Q32 W3Q40		0.70				0.75		
PC4:perceptions of safety	Perceived safety when there is a driver/steward on the DS	W1Q27 W2Q28 W3Q37				0.94				0.96

As shown in Table 5, four new components were classified in the PCA model. The first component identified perceived usefulness or satisfaction, reflected how useful or satisfying the driverless shuttle system was evaluated by respondents. The second component clustered, measuring the overall comfort perception and onboard convenience statement. The third component consists of three items, which compared the perceived travel time and perceived speed of driverless shuttle with other travel modes (e.g. car and regular bus) for the same travel distance. The last component with a single item investigated how safe respondents feel about the driverless shuttle. In total, four components explained 0.87 and 0.91 variances of information collected from two waves, which reached the recommended value of 0.50. Consequently, factor analysis roughly grouped the items and reduced the dimensions of tanglesome data. The second, third and fourth components were integrated into the hierarchical service quality framework influencing behavior intentions. Furthermore, the first component, perceived usefulness or satisfaction, was treated as a moderator between the service quality structure and behavior intentions.

[this page is supposed to be blank]

4. Methodology

The aim of this study is to enrich the longitudinal analysis of the public's evaluation and acceptance of the driverless shuttle (DS) service on a regular basis. To obtain an in-depth understanding of the longitudinal changes in people's attitudes and perceptions of the DS service over time, we first conduct the latent growth curve model (LGCM) analysis. Subsequently, to uncover the temporal pattern of longitudinal user acceptance of the DS service, a longitudinal structural equation model (LSEM) was developed.

4.1 Capturing Dynamics of Service Evaluation Changes through Latent Curve Analysis

One of the objectives of this thesis is to understand the longitudinal changes in people's evaluations of the DS service while exploring the influences of social-demographic, individual characteristics and travel attributes on them. The traditional analytical methods predict longitudinal change by using parameters that are common across cases. In other words, they smooth up the repeated measure to obtain a mean latent trajectory that cannot be directly observed for the entire population. Nonetheless, underlying individual time trajectories could not be captured by a simple regressive model. For this reason, a different approach to analyze longitudinal change is required. One of the newer analytical methods, latent growth curve models (LGCM), was applied to the current study. LGCM refers to "the statistical methods that allow for the estimation of inter-individual variability in intra-individual patterns of change over time" (Bollen and Curran, 2006). Moreover, differing from other approaches, LGCM can flexibly deal with various complex issues, such as discretely scaled repeated measures, unequally spaced time points and the inclusion of time-varying covariates (Curran et al., 2010).

LGCM can be specified as unconditional or conditional models. Unconditional models analyze the latent growth trajectories without any contribution of other variables, while conditional models take account of covariates for explaining the individual diversity lies in the general trend of the entire population. In the following sections, this method from Bollen and Curran (2006) will be recalled. In the LGCM context, "items" in the questionnaire is called "indicators", and "components" is equivalent to "repeated measures".

4.1.1 Capturing the Average Intra-individual Patterns of Service Evaluation Changes through Unconditional Model

The unconditional model is a powerful model for estimating the average trend of intra-individual growth patterns across the entire sample. The unconditional model only includes repeated measures that can be measured by directly observed indicators, analyzing the patterns within individuals without any contribution of other variables. The unconditional model aims to recognize the trajectory for the entire sample at first. It helps to decide whether the conditional model is necessary to be applied by checking individual trajectories around the mean trajectory.

The unconditional LGCM can be structured by a two-level model. The first level model gave the equation (1) of individual trajectory y_{it} at time point t for individual i. y_{it} is described by the intercept of the individual trajectory α_i and slope of the individual trajectory β_i along with trend variable λ_t. Trend variable λ_t captured time interval between waves of the panel survey. Additionally, disturbance term ε_{it} is attached to the end of the equation with the assumption that $E(\varepsilon_{it}) = 0$ for all t and i. Next, at the second level as shown in Equation (2) and Equation (3), intercept and slope are composed of the fixed components μ_α, μ_β and random components $\zeta_{\alpha i}$, $\zeta_{\beta i}$. Mean intercept μ_α and mean slope μ_β describe the central tendencies of intra-individual patterns of all cases in the sample. Namely, $\mu_\alpha = E(\alpha_i)$, $\mu_\beta = E(\beta_i)$. Disturbance $\zeta_{\alpha i}$ and $\zeta_{\beta i}$ reveal the inter-individual deviation from the central tendencies, where $E(\zeta_{\alpha i}) = 0$, $E(\zeta_{\beta i}) = 0$ are assumed for all i. Note that $VAR(\alpha_i) = VAR(\zeta_{\alpha i}) = \Psi_{\alpha\alpha}$, $VAR(\beta_i) = VAR(\zeta_{\beta i}) = \Psi_{\beta\beta}$ are only satisfied in the unconditional model. The higher values of $\Psi_{\alpha\alpha}$ and $\Psi_{\beta\beta}$, the larger inter-individual diversity lies in the trajectories of all individuals. By integrating the second level model to the first one, the reduced form is obtained by Equation (4) to depict an overall picture with all terms determining an individual trajectory.

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (1)$$
$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \quad (2)$$
$$\beta_i = \mu_\beta + \zeta_{\beta i} \quad (3)$$
$$y_{it} = \left(\alpha_i + \lambda_t \mu_\beta\right) + (\zeta_{\alpha i} + \lambda_t \zeta_{\beta i} + \varepsilon_{it}) \quad (4)$$

Where, $i = 1,2,3 \ldots I$, *I equals to the total number of individuals in the sample;* $t = 1,2,3 \ldots T, T$ *is the total time points (waves) of the survey period.*

In addition to the aforementioned assumptions of means [i.e. $E(\varepsilon_{it}) = 0, E(\zeta_{\alpha i}) = 0,\ E\left(\zeta_{\beta i}\right) = 0$], correlations between trajectory parameters (i.e.α_i, β_i) and disturbance term ε_{it} are deemed to be zero.

4.1.2 Capturing Inter-individual Variability in the General Service Evaluation Changing Trend through a Conditional Model

Based on unconditional model expected results, the existence of distinct trajectories for each case should have been recognized. Suppose the variance of intercept and slope in the unconditional model is statistically significant. In that case, it refers that the differences in the initial values and the rate of change over time between individuals need to be further explained. Hence, a conditional model is suggested to be applied to the data.

The conditional model involves moderators in the form of covariates. Specifically, covariates are separated into time-invariant covariates (TICs) and time-varying covariates (TVCs). Adding TICs allows for exploring which factors contributing to the individual variability around the mean trajectory. While TVCs have different values at each time point, influencing indicators accordingly.

Time-invariant covariates refer to moderators that do not change over time and only influence random intercepts and random slopes. The addition of TICs also makes it possible to predict which individuals have relatively higher initial value and which have relatively steeper growth rate in the sample. In the conditional model with TICs, the model at the first level is the same as the unconditional model but differ at the second level. As displayed in Equation (5) and Equation (6), TICs are noted by x_n (n = 1,2,3, … , N) and are added to the second level of the model. γ_α and γ_β provide the expected difference in the individual trajectories for one unit change in the TIC. Consequently, one unit change in the

x_n results in γ_α change of intercept and γ_β change of slope, compared with baseline case.

$$\alpha_i = \mu_\alpha + \sum_n \gamma_{\alpha n} x_n + \zeta_{\alpha i} \quad (5)$$
$$\beta_i = \mu_\beta + \sum_n \gamma_{\beta n} x_n + \zeta_{\beta i} \quad (6)$$

Where $n = 1,2,3,\dots,N$, N *equals to the number of TICs included in the model.*

TVC changes its value over time and has significant correlations with repeated measures at every time point. Unlike TIC, which addresses the adjustment of the model at the second level, TVC is conceptualized at the first level as Equation (7). The influence of ω_t on y_{it} was firstly removed, then the growth process on adjusted y_{it} is conducted. In other words, after deducting the influence of TVC, the remaining part of repeated measures is analyzed.

$$y_{it} = \alpha_i + \lambda_t \beta_i + \gamma_t \omega_t + \varepsilon_{it} \quad (7)$$

Where $t = 1,2,3 \dots T, T$ *is the total time points (waves) of the survey period.*

At last, simultaneous inclusion of TIC and TVC will change the model into the integrated form, which combines the first level model with TVC as Equation (7) and the second level model with TICs as Equation (5) and Equation (6).

4.1.3 Variable Selection

As the items assessing the DS service quality are likely to be closely correlated, and multicollinearity may lead to biased, a principal components analysis was conducted in Section 3.3. During the analysis, corresponding principal components and their items were assigned to three service quality attributes (PC2: perceptions of comfort and convenience, PC3: perceptions of travel time and speed, PC4: perceptions of safety). However, as shown in Section 3.3 Table 5, only comparative speed, comparative travel time, perceived safety, perceived general comfort and perceived convenience on board were collected in all three waves. However, LGCM requires at least three repeated measures for the regression of latent curves. Therefore, only functional attributes and safety attributes were analyzed as complete constructs.

An attempt was initially made to utilize second-order LGCM for the analysis of functional attributes. The advantage of second-order LGCM is that it can contain more information by keeping the disturbances of multiple indicators. Nevertheless, the trajectory parameters at the second level were not satisfactory due to the complexity of the model. Thus, a method to convert a latent construct with multiple indicators into a single repeated measure is needed. The most common way is to achieve this by calculating the factor scores. Ultimately, functional attributes were converted to a directly observed variable (also called single repeated measure) 'Functional Attributes (FA)', enabling a standard LGCM to be performed. Although details of each observed indicator were partially lost, the model parameters became significant. That is, the compression of information helps the model to interpret the data successfully. Safety attribute was measured only by a single item, whereas it was kept in the model as 'Safety Attribute' to ensure the completeness of the hierarchy of transit needs. Additionally, the hedonic attribute was considered not possible to address latent curve analysis due to the lack of observed data for one out of three items at the first time point. Hence, instead of the complete structure, two items, convenience on-board and general comfort, were analyzed to reflect the longitudinal change in the hedonic attribute.

In conditional models, moderators were divided into time-invariant covariates and time-varying covariates. According to the conceptual framework of the multi-level model on automated vehicle acceptance (MAVA) coined by (Nordhoff, et al., 2019), TICs and TVCs were also selected. The descriptive statistics of these variables were summarized in Section 3.3.

Socio-demographic variables including gender, age, education level, employment status, car ownerships were deemed as TICs influencing the inter-individual differences. In addition, travel attributes and individual personality were taken into account in the conditional model as well. Individual personality in terms of technology savviness was measured in detail by four items interviewed in the third wave. We believe that this feature did not change with time in the short term, so it was entered into the model as one of the TICs. Meanwhile, the familiarity to underlying technology supporting driverless shuttle was included as TIC.

Pertaining to TVCs, travel attributes in this study were reflected in the weekly use frequency of public transport as a connecting mode to the transit hub. The variables notation 'PUT' was given. We observed that the value of this variable changed over time at three time points. Thus, 'PUT' was treated as TVC. Another TVC involved in the conditional model was measured by last week usage of driverless shuttle service, noted by 'LWuse'. 'LWuse' implied the degree of current user activity and can be considered a moderator to explain the inter-individual difference in attitude change.

4.1.4 Model Specification

In order to interpret the panel data properly, the model specification was built. In the diagram, a standardized notation of SEM was adopted. Rectangles represent observed variables. Ellipses represent unobserved variables such as latent factors and residuals. The straight single-headed arrows imply the factor loadings or regression coefficient, while curved double-headed arrows indicate the covariance between variables.

As an example, a general specification path diagram for the unconditional model is displayed in Figure 8. y_t $(t = 1, 2, 3)$ represents the repeated measures of the target variable at three interview time points. In the following chapter, y_t refers to functional attribute, safety attribute, general comfort and convenience on board, respectively. In order to obtain both fixed and random components of intercept by estimating the mean and variance of α_i, intercepts of all three y_t were set to zero. Meanwhile, the factor loadings from intercept α_i to y_1, y_2, y_3 were all fixed to values of 1. Namely α_i has equal impacts on all the repeated measures across three waves. Disturbance at each time point e_t $(t = 1, 2, 3)$ was attached to corresponding repeated measure by setting regression weight to one. The mean values of disturbances were assumed to be zero.

Time coding is also an essential specification of the models. Since the survey scheme was not strictly designed with equal time interval, loadings of slope factor β_i should be modified from the standard coding. It is worth recalling that the survey dates of three waves were in March, September and December. Consequently, time intervals between waves were calculated as six months and three months, resulting in cumulative time series: starting month, the sixth month and the ninth month. In LGCM, three time points can only adopt linear regression. Therefore, time coding 0, 2, 3 was assigned to trend parameters λ_1, λ_2, λ_3. That is, factor loadings from slope β_i to each latent variable were set to 0, 2, 3 accordingly. $\lambda_1 = 0$ was set so that the mean of repeated measures at the initial time point can be represented by $E(\alpha_i)$. In addition, intercept α_i and slope β_i were allowed to be correlated by adding a double-headed arrow on the disturbances u_1 and u_2 of them. Furthermore, disturbances of intercept and slope were also set to have a mean of zero.

Figure 9 shows the path diagram of conditional model specification based on the unconditional model, which involves TVC and TIC simultaneously. ω_1, ω_2, ω_3 represent TVC at each time point, influencing the repeated measure y_1, y_2, y_3 correspondingly. x_1 illustrates one of the TICs, explaining growth curve parameters by pointing to intercept factor α_i and slope factor β_i. Note that, TVC values at each time point were allowed to be correlated with each other and with the TIC in the model.

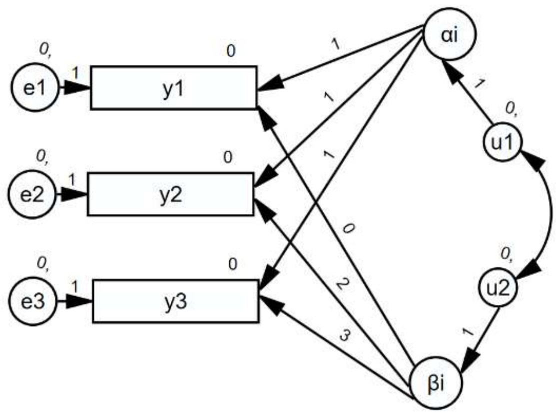


Figure 8 Unconditional Model Specification

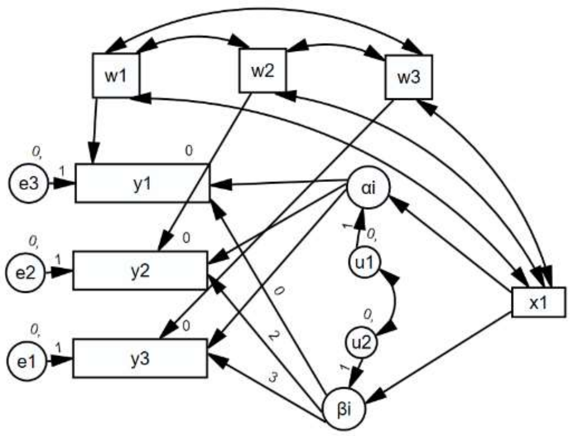


Figure 9 Conditional Model Specification

4.2 Capturing Temporal Pattern of User Acceptance through Longitudinal Structural Equation Modeling

In the abovementioned section, LGCM was constructed and estimated from a structural equation models perspective. Hence, as the basis of the estimation method of LGCM, the concept of structural equation modeling (SEM) is necessary to be introduced. Moreover, after analyzing longitudinal changes in service quality attributes, the causal relationships between BI and the service quality attributes need to be verified. For this reason, SEM stated in *Structural Equations with Latent Variables* (Bollen, 1989) will be recapped in this section.

4.2.1 User Acceptance Model Framework

In combination with the existing user acceptance models for driverless shuttle mentioned before in the literature review and the questionnaire design for the panel survey, BI and WOM were chosen as dependent variables of the acceptance models in this study. Conceptual frameworks of models were proposed for adopters and non-adopters separately in Figure 10 and Figure 11

4.2.1.1 Using Intention and the Word-of-mouth Recommendation

For the successful promotion of driverless shuttle service, actual usage of the new transportation mode as the daily travel pattern should be the ideal research target. However, due to the short trial period and the setting of response scales in the questionnaire, the changes in cumulative use frequency ranged by “never use, 1-5 times, 6-10 times, 10-15 times, more than 15 times” cannot be observed between waves. Alternatively, the survey question “last week use of the driverless shuttle service” was attempted to be included in the model as the dependent variable in the longitudinal study. Instead of actual usage, using intention was the dependent variable affected by explanatory variables in the model. Using intention, also named behavior intention (BI), has been proved as a critical and well-established predictor of system usage in Section 2.3. Simultaneously, the vital role of using intention was also advocated by many studies in the context of user acceptance of driverless shuttle, as reviewed in Section 2.3 Table 1. In the previous study, determinants of BI were proposed as attitude, social norms and needs by Guo et al. (2021). While the longitudinal model for the project in Kista reveals that BI was influenced by safety, ride comfort and travel time reliability (Chee et al., 2021). In the current study, BI was indirectly determined by three service quality attributes through mediators, affecting WOM in sequence.

Guo et al. (2021) introduced word-of-mouth recommendation (WOM) into the acceptance process of driverless shuttle and explored the contributing factors of WOM. The model results demonstrated that attitude, social norms and needs positively influenced WOM and BI at the same time. However, the relationship between BI and WOM was considered insignificant (H6). Therefore, this paper also attached WOM to the model as the dependent variable secondary to BI. The purpose was to try to strengthen the understanding of WOM's role in the driverless shuttle context.

As one of the most prominent behavioral outcomes, WOM was believed to be positively influenced by the level of perceived service quality (Harrison-Walker, 2001). Thereupon, the hypothesis was proposed that perceptions of service quality indirectly determine WOM after mediated by satisfaction (H5). In the same way, this path was also imposed on the non-adopters model. Hence, the positive impact on BI of usefulness was examined (H5'). Moreover, in order to further discuss the significance level of the correlation between WOM and BI, hypothesis 6 (H6) was tested separately for adopters and non-adopters.

As aforementioned, three explanatory variables, two mediating variables and two dependent variables were included in two acceptance models. The existence of further explanatory variables cannot be denied, which need to be considered within the driverless shuttle context. However, based on the collected data, we believe that the significant determinants of this trial were addressed in the current study.

4.2.1.2 Satisfaction and Perceived Usefulness of the Shuttle Service

In the technology acceptance model (TAM), perceived usefulness was coined by (Davis, 1989) as "the degree to which a person believes that using a particular system would enhance his or her job performance". Notably, perceived usefulness was identified as the primary determinant of intention in TAM. After Davis, Unified theory of acceptance and use of technology (UTAUT; Venkatesh et al., 2003) collaborated TAM with other seven mainstream acceptance models. It integrated perceived usefulness into a general construct named performance expectancy (PE). Afterwards, abundant studies have repeatedly verified PE as the essential determinants of intention to take driverless shuttles (literature reviewed in Section 2.3 Table 1).

Under the same project, the prior intention study for a single time point was based on combined TAM and TPB (C-TAM-TPB) model structure, regressed attitudes, social norms and needs (equivalent to the term “perceived usefulness”) constructs on BI. On the other hand, the current longitudinal study adopted the service quality evaluation framework and TAM as the supporting theories. Perceived usefulness was extracted from TAM as a direct determinant, influencing BI and WOM jointly with the indirect effects of service quality characteristics.

As early as in TAM, Davis pointed out that the relationship between ease of use and usefulness might be rather causal than parallel. According to this argument, the relationship between service quality attributes, perceived usefulness and BI was also proposed. It was proposed that instead of parallel structure, perceptions of service quality attributes could be a causal antecedent to perceived usefulness, and through perceived usefulness, indirectly affect BI. In other words, similar to the role of satisfaction, perceived usefulness was regarded as a mediator to study the determinants of BI of non-adopters.

4.2.1.3 Quality of the Driving Shuttles Service

Hitherto, most of the current works aimed to explain how using intention was directly affected by several determinants. Specifically, both the model based on the same data set as this study (Guo et al., 2021) and the model with the most similar determinants (Chee et al., 2021) adopted this structure.

In Section 3.4, four principal components were obtained by dimensionality reduction through factor analysis (PC1: perceived usefulness or satisfaction, PC2: perceptions of comfort and convenience, PC3: perceptions of travel time and speed, PC4: perceptions of safety). In the beginning, direct regression of the components on BI was also conducted. Nevertheless, the parallel and direct effects of these four principal components on BI did not properly fit the data in the longitudinal model when extending to two time points.

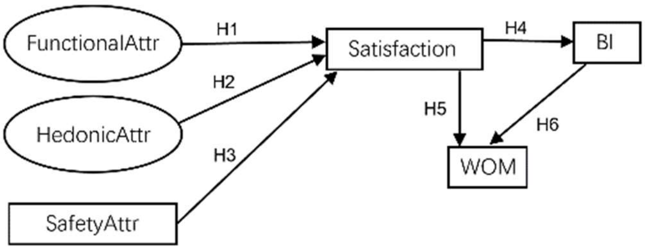


Figure 10 Conceptual Framework for Adopters

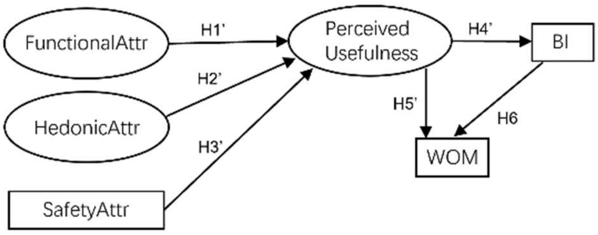


Figure 11 Conceptual Framework for Non-adopters

The correlation between service characteristics and acceptance of ADS-AVs was accentuated by Nordhoff et al. (Nordhoff et al., 2016). In addition, referencing from the adaption of the hierarchy of transit needs for service quality attributes (Chee et al., 2020), PC2 to PC4 were exactly in accord with the three levels in Section 2.5 Figure 3. Therefore, the concept of public transit service quality was introduced into the model. Subsequently, PC2 to PC4 were renamed by “functional attributes (FA)”, “safety attributes (SA)” and “hedonic attributes (HA)”, respectively. Similarly, it has been verified that the mechanisms of these three attributes functioning on satisfaction were in line with basic, performance and exciting factors of the three-factor theory (also mentioned in Section 2.5).

For driverless shuttle users, safety attributes were considered the most critical factor. No matter the evaluation on safety attributes was low or high, it significantly affected the overall evaluation of service quality (H3). Functional attributes reduced the comprehensive evaluation of service quality when it was in low performance (H1). However, once the functional attributes (located at the bottom of the triangle) were considered satisfactory, they became irrelevant to the overall level of service quality. At the top of the triangle, hedonic attributes dominated service satisfaction only in a high-performance situation (H2). Hedonic attributes were deemed to have no significant contribution to the overall service evaluation when their performance was perceived as low. Above all, three service quality attributes were hypothesized to determine whether the respondents were satisfied with the driverless shuttle service (H1, H2, H3).

Question "satisfaction of the driverless shuttle ridden trip" in PC1 measured the satisfaction degree with a driverless shuttle trip among experienced users (adopters). The remaining items in PC1 were aggregated as perceived usefulness of inexperienced respondents (non-adopters). Thus, PC1 was decomposed into two subfactors, satisfaction and perceived usefulness, as the dependent variables of three service quality attributes for adopters and non-adopters, respectively.

As three hypotheses were proposed for adopters, likewise, three corresponding assumptions were presumed for non-adopters. The slight difference was that usefulness substituted satisfaction due to the lack of ridden experience, and consequently, H1', H2' and H3' were obtained. The mediator effect of satisfaction between service quality and behavior has been discussed in Section 2.5. In consequence, service quality was assumed to have both direct effects on satisfaction and indirect effects on BI after mediated by satisfaction (H4).

4.2.2 Longitudinal Structural Equation Model

The longitudinal structural equation model (LSEM) is developed in the form of structural equation modeling (SEM) to analyze the changes in the judging criteria regarding DS service adoption among respondents and evaluate the sequential effect of prior propensity on respondents' continued behavior intentions.

In essence, the notion of SEM is not a single statistical technique but an integration of various multivariate techniques under one model fitting framework. It collaborates the measurement theory, factor analysis, simultaneous equations and multiple regression all together. Instead of one dependent variable and a set of predictors of it in the traditional method, SEM treats the relationships between variables as a system that allows for the involvement of constructs with measurement errors. In general, SEM can be deemed as a causal model with latent variables. Latent variables are the unobservable variable hidden behind the observed items, reflecting the common cause of them. As expressed by Equation (8), observed score X is comprised by true score t and error term e. To separate true score from the systematic and measurement errors, we need multiple observed items to build identified equations. Then the loading weights from latent variable to observed items are obtained.

$$X = t + e \tag{8}$$

Concerning the SEM model construct, it consists of the measurement model and the structural regression model. Accordingly, confirmatory factor analysis (CFA) is applied to the measurement model, and causal relationships will be tested in the structural regression (SR) analysis. The covariance matrix between factors will be estimated by CFA and then become the input of SR to obtain the regression coefficients between factors.

Regarding the input and output of SEM, the input of SEM could be a set of qualitative causal hypotheses derived from existing theories or empirical study results. At the same time, the outputs of SEM are numeric estimates for hypothesized effects and how supportive the data are for the implications of the model (Pearl, 2012; Kline, 2016).

Identification

In general, identification can be verified by checking the number of known and unknown parameters in the model. If the number of unknown parameters does not exceed the number of unknown parameters, the unique solution of equations can be found. In particular, The model identification of SEM referenced the two-step rule stated by (Bollen, 1989, pp. 328). The first step treats the measurement part as a CFA model and confirms that the respecified CFA model is identified. Next, the second step treats the SR model as a path analysis model with observed variables and examines if it is identified. The model that can meet the requirements of both steps is deemed as identified. In this study, models depicted in Figure 12 and Figure 13 were all proved to be identified.

4.2.2.1 Confirmatory Factor Analysis

In contrast to explanatory factor analysis (EFA), confirmatory factor analysis (CFA) requires a priori construct specification presumed by the researcher. The purpose of CFA is to confirm the proposed hypotheses of causal effects with the sample data. It is noteworthy that the results obtained from CFA need a guarantee of both reliability and validity. For this reason, reliability and validity tests play an essential role in the measurement model estimation.

Reliability and validity test

Internal reliability of the model refers to the preciseness of observed scores in the sample. The internal consistency of the measurement model was examined by composite reliability (CR). CR could be understood as the ratio of total true score variance to the total observed score variance (Brunner and SÜβ, 2005). When facing different scales or different scale types of the items, CR has been considered a more robust way than Cronbach's alpha (Netemeyer et al., 2003). In the current study, except items measuring attitude towards technology were in a seven-point Likert scale, all the other items were in the same five-point Likert scale. Therefore, CR was reported to evaluate the consistency of responses across the items under the same latent variable. The thresholds for CR depend on the numbers of observed items. A threshold of reasonable value could be from 0.60 and up (Bagozzi and Yi, 1988). The fewer items designated to a construct, the lower the reliability level will be.

The validity of the model was verified by convergent validity and discriminant validity. Convergent validity supports the appreciable magnitude of intercorrelations between items measuring the same latent variable. The average variance extracted (AVE) was the measure to check convergent validity. AVE measures the amount of construct captured variance in relation to the amount of measurement error variance. When AVE above the criterion value of 0.50, the convergent validity can be verified. On the opposite, discriminant validity implies that the correlation between items belonging to different latent variables is relatively low or close enough to zero. When all the inter-construct correlations are less than the root square of AVEs, no overlap between different constructs is suggested (F. Hair Jr et al., 2014). As a result of the examination, discriminant validity can be proved.

4.2.2.2 Structural Equation Model Analysis

SEM analyzes the observed covariance matrix instead of raw data. Then this observed covariance matrix is simplified by the underlying structure that is specified in SEM. The specified structure will yield an implied covariance matrix which can be compared to the observed covariance matrix. Subsequently, when the model is identifiable, unknown model parameters can be estimated. Instead of focusing on the difference between observed and predicted individual values, SEM estimation aims to minimize the discrepancy between the observed covariance matrix and the model implied covariance matrix while calculating the unknown parameters. More specifically, if the identifiable model is correct, the fundamental hypothesis of SEM expressed in Equation (9) will fail to be rejected (Lawley, 1940, pp. 1-2). S represents sample variance and covariance matrix, while $\hat{\Sigma}(\theta)$ represents the implied covariance matrix of estimated parameters θ. In turn, as one of the important model parameters, the regression coefficients reflecting correlations between variables will provide evidence to examine the causal hypotheses.

$$\text{H0: } S = \hat{\Sigma}(\theta) \qquad (9)$$

In the case of SEM, the maximum likelihood (ML) technique is prevalently adopted as an estimator for continuous data. As is known to all, ML estimates model parameters that can maximize the likelihood function of sample data. The properties of asymptotically unbiasedness and efficiency make ML prevailing and desirable. Theoretically, Likert scales are ordinal because of their ordered and categorized response options. Whereas, it has been suggested that variable with five or more than five categories can be deemed as approximately continuous in the analysis (Norman, 2010). The sample data of the current project were measured in a five-point Likert scale and seven-point Likert scale, which meets the abovementioned standard. Therefore, ML estimation was adopted in the current study.

Normality and Multicollinearity

However, one of the essential assumptions of ML is that the data come from a multivariate normality distribution. (Lawley, 1940; Bollen, 1989, pp. 131-135) Hence, the normality test was conducted. The skewness and kurtosis values for each variable were used to evaluate the normality of each independent variable. In Appendix A, the test results are displayed. All the absolute values of skewness and kurtosis were both less than 2. Meanwhile, Mardia's index of multivariate kurtosis were both larger than the indicative value of 5. Therefore, the test indicated that all the independent variables yield univariate normal distribution, but multivariate distribution significantly departs from expected normality. As the alternative solution to non-normal multivariate distribution, the Bollen–Stine bootstrap was applied with 2000 samples to compensate for the non-normal multivariate distribution of the data.

Another critical assumption of LM is that the independent variables of the hypothesis model do not have multicollinearity among them. The multicollinearity was detected among independent variables in the regression model. According to the diagnosis results, tolerances were all larger than 0.2 and variance inflation factors (VIF) were all less than 5, indicating no multicollinearity among the independent variables. One illustrative output is shown in Appendix B.

Assessing Model Fit

In the SEM context, a set of model fit indices were suggested to be reported at a minimum (Kline, 2016, pp. 269). According to the list and other commonly used indices, model fit statistics were selected. Model chi-square χ^2_{M} with its freedom df test the null hypothesis of SEM in Equation (9). Goodness of fit index (GFI) measures the relative amount of the covariances and variances in the observed matrix that are predicted by model implied matrix. Root Mean Square Error of Approximation (RMSEA) is considered an absolute fit index, measuring the degree to which the hypothesized model departs from a perfect mode. Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) are incremental fit indices that compare the fit of the hypothesized model against the null model. In respect to the evaluation of the values of these indices, "rules of thumb" cutoff criteria are reviewed here. In practice, χ^2_{M} / df less than 5 is considered acceptable (Marsh and Hocevar, 1985; Awang, repr. 2013, 2012). The cutoff values of 0.90 for CFI and TLI and of 0.05 for RMSEA were suggested as satisfactory fit (Awang, repr. 2013, 2012). Additionally, values of 0.90 or greater for GFI indicate an acceptable model fit (Pituch and Stevens, 2016). However, some literature advocated that GFI value of 0.8 or more is also an acceptable fit (Doll et al., 1994; Baumgartner and Homburg, 1996; Cheng, 2011; Awang, repr.

2013, 2012). Moreover, values of RMSEA between up to 0.08 or 0.10 are considered acceptable (Browne and Cudeck, 1992; Hu and Bentler, 1999).

4.2.2.3 Two-time-point Longitudinal Structural Equation Model Specification

Due to the reason that 'Satisfaction' was not available to the non-adopters response group, two different conceptual frameworks were proposed respectively for adopters and non-adopters (as mentioned in Section 4.2.1). In this section, the adopters' model (Figure 12) and non-adopters model (Figure 13) are specified based on the previous conceptual framework.

As the first step of specification, a common measurement model for both adopters' and non-adopters was specified. Based on the results from components analysis in Section 3.4, three service attributes, 'Functional Attributes', 'Hedonic Attributes' and 'Safety Attribute', were set as explanatory variables. The first variable, 'Functional Attributes', was measured by three indicators regarding speed and travel time compared to regular bus and car. Then, 'Hedonic Attributes' was measured by three indicators about comfort and convenience perceived by the respondents. They are both latent variables with multiple indicators. However, the third variable, 'Safety Attribute', was measured with a single indicator in respect to the on-board steward.

Subsequently, the SR model for a single time point was build according to the causal hypotheses proposed in Section 4.2.1. 'Functional Attributes', 'Hedonic Attributes' and 'Safety Attribute' were the determinants of mediating variables. Then, through mediating effects, dependent variables were affected by the determinants indirectly. In the adopters model, 'Satisfaction' is the mediating variable measured by how satisfied the respondents were with the riding experience. While in the non-adopters model, the role of mediator was played by latent variable 'Usefulness', which was measured by three indicators evaluating the degree to which respondents; travel needs can be met by driverless shuttle. For both models, influences behavior intention and word-of-mouth recommendation were the dependent variables influenced by 'Satisfaction' or 'Usefulness'. Additionally, BI also have impacts on WOM.

Afterwards, the SR model for two time points was established by involving variables in two waves. The connections between two time points were performed in the model by attaching intertemporal paths to explanatory variables, mediator variables, dependent variables and moderator variables. Equation (10) illustrates the variable at the previous time point $t-1$, while Equation (11) expresses the variable at subsequent time point t. The intertemporal effect is added to the dependent variable z_t in the form of cz_{t-1}, where c stands for intertemporal coefficient. y_{t-1}, y_t represent determinants of z_{t-1}, z_t; while a, b are the regression coefficients andδ_{t-1}, δ_t represent the disturbance. Consequently, the perception and evaluation from the last time point can be mapped to that at present by $z_{t-1} \rightarrow z_t$. Importantly, to solve the issue of dependence of error variance caused by the same indicators across time, we allowed disturbances that belong to the same indicator to correlated with each other.

$$z_{t-1} = a + by_{t-1} + \delta_{t-1} \quad (10)$$

$$z_t = a + by_t + cz_{t-1} + \delta_t \quad (11)$$

To make sure the model is identifiable, we imposed constraints on model parameters. All the disturbance terms were fixed to means of zero and loading weights of one, which is usually called unit loading identification (ULI) constraint (Kline, 2016, pp. 148). By setting the three items as marking variables in turns, we compared the difference of factor loadings between the two waves when each item was set as marking. The item with the smallest difference of loadings was finally selected. As a result, marker variables were set to W2Q31 and W3Q39 for 'Hedonic Attributes' while W2Q36 and W3Q44 for 'Functional Attributes'.

Additionally, in order to have more clue on how to encourage potential users to adopt the driverless shuttle service in the future, we included two moderator variables in the non-adopters model based on the results from the LGCM section. The time-varying moderator 'PUT' was assumed to have direct effects on contemporaneous BI. The degree of passion for innovative technology (attitudes towards technology) was deemed time-invariant, indirectly influencing BI and continuance intention through perceived usefulness.

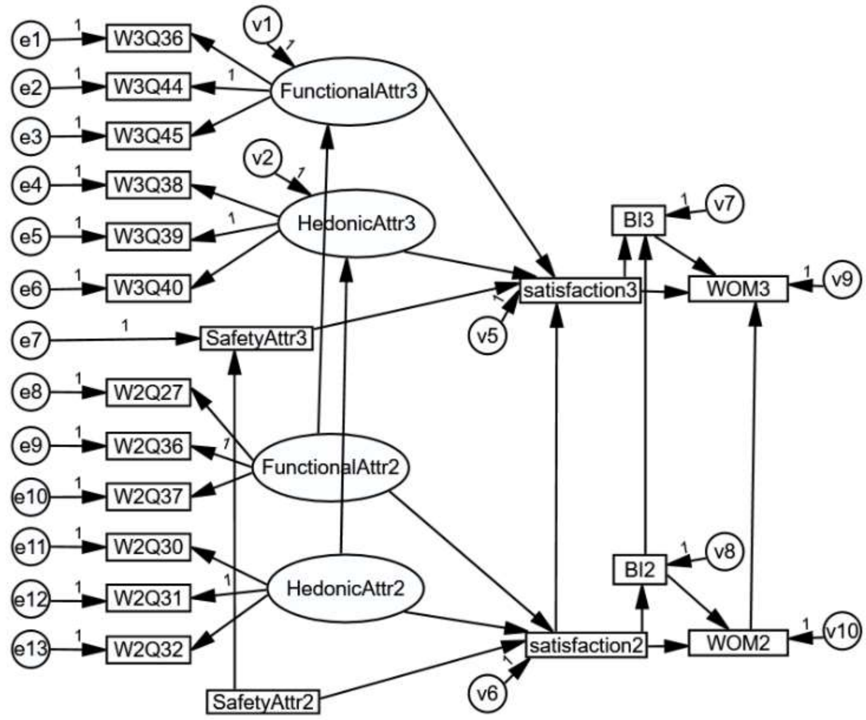


Figure 12 Two-time-point User Acceptance Model Specification for Adopters

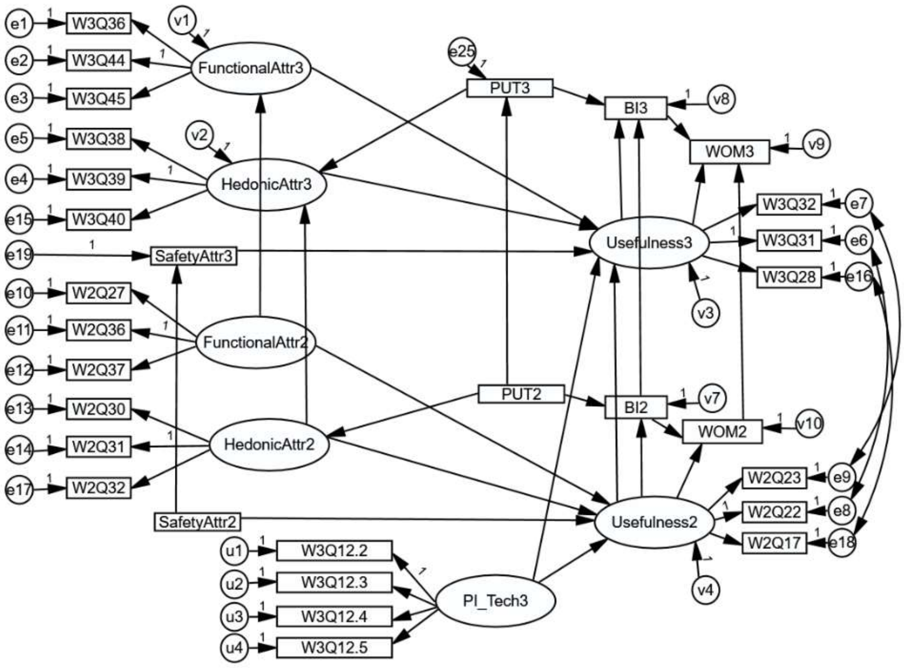


Figure 13 Two-time-point User Acceptance Model Specification for Non-adopters

[this page is supposed to be blank]

5. Results

Results of the latent growth curve model (LGCM) and the longitudinal structural equation models (LSEM) are presented in this chapter. The average growth trajectory of service evaluations and interindividual variability in the general trend will be revealed. Further, hypotheses proposed in the conceptual model for adopters and non-adopters will be examined. Meanwhile, the influences of different socio-economic profile and various individual characteristics on the evaluations of the DS service will be analyzed.

5.1 Dynamics of Service Evaluation Changes Towards the Driverless Shuttles

In this section, a latent growth curve model was employed to understand the changes in the respondents' perceptions of the DS service over time and explain how socio-economic profile and individual characteristics influence the changes. It is noteworthy that the interest of analysis is the underlying estimated latent growth trajectories.

5.1.1. Unconditional Model

In unconditional models, the dynamic changes in 'Functional Attributes (FA)', 'Safety Attribute (SA)', 'Convenience', 'Comfort' were studied with respect to the trajectories for the entire sample and for each individual.

Before conducting standard LGCM, factor scores of the 'FA' construct at each time point were obtained from the original measurement model. In other words, three indicators of the construct were merged into composites that can be considered as a single repeated measure. The factor loadings of each indicator at the corresponding time point are displayed in Table 6. All the loadings are larger than the recommended value of 0.60, except for the loading of indicator 'W2Q27'. However, as the value of 0.561 is extremely close to 0.60, it is considered acceptable. Subsequently, the model reliability was tested by checking the composite reliability (CR) values. As shown in Table 6, all CR values exceed 0.70, resulting in the confirmation of the reliability of the measurement model. In the meanwhile, the criteria of average variance extracted (AVE) are met at all three points in time. All the values of AVE are beyond the cutoff value of 0.50, demonstrating the convergent validity of the measurement model. Last but not least, all the inter-construct correlations are listed in the lower-left corner of Table 7. All the correlations are less than the square roots of AVEs that are on the diagonal. Accordingly, the discriminant validity of the measurement model is tested. Ultimately, the composition of 'FA' is proved from the perspectives of both reliability and validity.

After obtaining the composites of 'FA', linear regression LGCM with three points in time was conducted. Trajectory parameters in four unconditional models were estimated separately. The estimated results are displayed in Table 8.

Table 6 Measurement Model Estimates of 'Functional Attributes' at Three Time Points

Construct	Indicator	Factor Loading	AVE	CR
FunctionalAttr1	W1Q23	0.634		
	W1Q33	0.790		
	W1Q34	0.699	0.505	0.752
FunctionalAttr2	W2Q27	0.561		
	W2Q36	0.797		
	W2Q37	0.829	0.546	0.778
FunctionalAttr3	W3Q36	0.634		
	W3Q44	0.839		
	W3Q45	0.753	0.558	0.789

W1Q23, W2Q27, W3Q36 corresponds to the question in relation to the perceived speed of a driverless shuttle compared to a normal bus.
W1Q33, W2Q36, W3Q44 corresponds to the question in relation to the perceived travel time of a driverless shuttle compared to a normal bus.
W1Q34, W2Q37, W3Q45 corresponds to the question in relation to the perceived travel time of a driverless shuttle compared to a private car.

Table 7 Correlation and Square Root of AVEs Matrix of 'Functional Attributes'

Correlation/Sqr(AVE)	FunctionalAttr1	FunctionalAttr2	FunctionalAttr3
FunctionalAttr1	0.711		
FunctionalAttr2	0.554	0.739	
FunctionalAttr3	0.484	0.647	0.747

Table 8 Unconditional LGCM Parameters Estimates

Functional	Estimate	Standard Error	P	Safety	Estimate	Standard Error	P
$\mu_{\alpha 1}$	1.589	0.030	0.000	$\mu_{\alpha 2}$	3.901	0.048	0.000
$\mu_{\beta 1}$	0.083	0.010	0.000	$\mu_{\beta 2}$	-0.077	0.021	0.000
$\Psi_{\alpha 1 \alpha 1}$	0.310	0.045	0.000	$\Psi_{\alpha 2 \alpha 2}$	0.391	0.134	0.004
$\Psi_{\beta 1 \beta 1}$	0.030	0.007	0.000	$\Psi_{\beta 2 \beta 2}$	0.054	0.023	0.017
$\Psi_{\alpha 1 \beta 1}$	-0.029	0.015	0.055	$\Psi_{\alpha 2 \beta 2}$	-0.089	0.052	0.085
Corr(α1,β1)	-0.301			Corr(α2,β2)	-0.615		
Convenience	Estimate	Standard Error	P	Comfort	Estimate	Standard Error	P
$\mu_{\alpha 3}$	3.053	0.039	0.000	$\mu_{\alpha 4}$	3.535	0.038	0.000
$\mu_{\beta 3}$	0.015	0.015	0.322	$\mu_{\beta 4}$	-0.027	0.016	0.09
$\Psi_{\alpha 3 \alpha 3}$	0.265	0.073	0.000	$\Psi_{\alpha 4 \alpha 4}$	0.317	0.086	0.000
$\Psi_{\beta 3 \beta 3}$	0.020	0.013	0.114	$\Psi_{\beta 4 \beta 4}$	0.035	0.014	0.015
$\Psi_{\alpha 3 \beta 3}$	-0.030	0.028	0.284	$\Psi_{\alpha 4 \beta 4}$	-0.055	0.033	0.093
Corr(α3,β3)	-0.419			Corr(α4,β4)	-0.522		

Dynamic Changes of Functional Evaluations

First of all, the intercept and slope of the mean trajectory for 'FA' were estimated to be $\hat{\mu}_{\alpha 1}$=1.589, $\hat{\mu}_{\beta 1}$=0.083, respectively. Both of them differed significantly from zero, indicating that the intraindividual pattern of 'FA' is an increasing line on average. The results imply that respondents initially evaluated the speed and travel time of DS service unsatisfactory when compared the service to regular bus and car. Nevertheless, respondents' evaluations were steadily increasing after the first interview. In other words, the basic needs of a public transport mode are not met by the driverless shuttle service at the present stage. Theoretically, this will significantly reduce the overall service satisfaction, but when the FA evaluation improves to a high-performance level in the future, the negative impact is expected to disappear. Therefore, the upward trend of FA should be maintained. This could be achieved by increasing operation speed and shortening travel time until a high functional performance of the driverless shuttle is met.

Dynamic Changes of Safety Concerns

The second attribute, safety attribute 'SA', was analyzed subsequently. It is shown in Table 8 that $\hat{\mu}_{\alpha 2}$=3.901 and $\hat{\mu}_{\beta 2}$=-0.077 significantly differ from zero, indicating that the average trend of intraindividual trajectories of 'SA' is also an increasing line. The finding reveals that the safety performance of the DS service was generally given a satisfactory evaluation, but showed a decreasing trend with the passage of time. It is possible to submit that people feel safe on the driverless shuttle when there is a steward on board, but the trust declined over the survey period. As a critical requirement of transit, the high-performance safety attribute of the service is supposed to promote overall satisfaction. However, there is a considerable risk that the evaluation from safe will slip to unsafe in the future, which will negatively influence the overall quality of service.

The given picture of safety concerns could be fuzzy because of the onboard steward's presence during the trial. Although the respondents generally perceived the driverless shuttle relatively safe in this trial, when the steward is removed in the future, the perception of safety performance might be severely impacted. Consequently, the low ratings of safety performance will give rise to a dramatic plunge in overall satisfaction. Although the development pattern of safety concerns was ambiguous, safety is still considered a critical determinant of service quality. Namely, safety would significantly affect the overall satisfaction regardless of the performance level. Operators could develop a steady transition from backing up by a steward to completely no human intervention on board. In addition, with auxiliary presentations of emergency plans for specific incidents, passengers' trust in the safety of driverless shuttle is expected to be enhanced. Eventually, the positive attitude of respondents towards the safety performance can be maintained as same as the present stage.

Dynamic Changes of Comfort and Convenience Concerns

Two items that substitute for the construct of hedonic attributes were noted by ‘Convenience’ and ‘Comfort’. The mean of initial evaluation concerning convenience on board was estimated as $\hat{\mu}_{\alpha3}$=3.053. Nevertheless, the mean slope of intraindividual trajectories of ‘Convenience’ did not significantly differ from zero, suggesting that respondents kept holding a moderate opinion on the perceived convenience on board in general. Another substitution is ‘Comfort’. the mean trajectory for ‘Comfort’ was portrayed by $\hat{\mu}_{\alpha4}$= 3.535, $\hat{\mu}_{\beta4}$= -0.027, indicating that the initial evaluation of general comfort was satisfactory on average but dropped over time.

The evaluations of comfort and convenience are generally higher than satisfactory level with a downward trend. Comfort and convenience as the hedonic requirements are fulfilled at the top layer of the needs’ hierarchy, so it should have a significant positive effect on the overall service satisfaction.

Therefore, the downward trend of hedonic attributes should be alleviated. The speed control pattern of the driverless vehicle could be optimized to moderate the discomfort caused by the acceleration and deceleration, and the overall comfort of the ride experience could also be improved by polishing up the design of on-board amenities. Such countermeasures are expected to turn the evaluation of comfort from decline to rise, and subsequently pursue the positive effects that hedonic attributes bring to overall service satisfaction.

Additionally, the interindividual variability around the intercept and slope of mean trajectory were revealed by the variance parameters of each variable, $\widehat{\Psi}_{\alpha i \alpha i}$ and $\widehat{\Psi}_{\beta i \beta i}$ (i = 1,2,3,4), respectively. The results demonstrate the existence of fluctuations around the mean trajectory resulted from interindividual diversity. Consequently, further analysis of interindividual differences is necessary to be conducted through the conditional model, and the incorporation of moderator variables is needed to better explain the variability in growth factors.

5.1.2. Conditional Model

To explore which factors contributed to the inter-individual variabilities around the mean trajectory pattern of the whole sample, social-demographics, individual characteristics, and travel attributes were involved as extra explanatory variables in the conditional models. The repeated measures were regressed on the time-varying variable (TVC), then the remaining random parts of intercept and slope were regressed on the time-invariant covariates (TICs).

Recall that TICs involved in the models are gender, education level, employment status, car ownerships, travel attributes, attitude towards technology and familiarity with underlying autonomous driving technology. Meanwhile, the selected TVCs are weekly use frequency of public transport connecting to the transit hub ('PUT') and last week use frequency of the driverless shuttles ('LWuse'). The coding and descriptive statistics of TVCs and TICs were summarized in Section 3.3 Table 2, Table 3 and Table 4. All the maximum likelihood estimates were displayed with their standard errors in Table 10 and Table 11. Additionally, according to the values of the chi-square test statistic, CFI, TLI and RMSEA, all conditional models are demonstrated to have excellent model fits.

5.1.2.1 Coefficients of Time-varying Covariates (TVCs)

Firstly, the estimated growth factors of 'PUT' in Table 9 reveal that public transport's average use frequency among the respondents is 1-2 days per week. The results indicate that most respondents did not frequently use public transport modes to connect to the transit hub initially. However, after the first interview, respondents gradually increased their use frequency of public transport to reach the nearby subway station or other transit modes over the survey period.

Subsequently, according to the regression estimates of 'PUT' on different aspects of service evaluations (as shown in Table 10), the increasing frequency of public transport usage is found to positively and significantly influenced the perceptions and evaluations of functional attributes, safety attribute and convenience in general. Although the predictive effects of public transport usage in the first and third wave were partially insignificant, the positive effects of public transport utilization are still considered non-negligible. Thus, it is possible to submit that the more frequently people use public transport (e.g. regular bus) as a feeder mode, the more satisfactory evaluations of the service they tend to give. Therefore, current public transport passengers are one of the potential user groups who are more likely to adopt the DS service. In this way, a special discount on SL ticket for the DS ride can be appealing to current regular bus passengers, leading to the shift from regular bus to driverless shuttles. Alternatively, decision-makers can collaborate the driverless shuttle with other available transportation modes in the monthly package of MaaS, providing seamless and integrated intermodal travel.

Table 9 Latent Growth Curve for Public Transportation Usage

PUT	Estimate	Standard Error	P
$\mu_{\alpha 5}$	2.419	0.060	***
$\mu_{\beta 5}$	0.108	0.019	***
$\Psi_{\alpha 5 \alpha 5}$	0.740	0.130	***
$\Psi_{\beta 5 \beta 5}$	0.04	0.021	0.054
$\Psi_{\alpha 5 \beta 5}$	-0.011	0.044	0.808
Corr(α5,β5)	-0.063		

(*** represents p-value<0.001)

Table 10 TVCs Coefficients Estimates (Standard Errors) for Conditional Models

	Predictor Variable	Coefficient Estimate	P
Functional Attributes	PUT1	**0.035(0.059)**	0.100
	PUT2	**0.046(0.064)**	0.025
	PUT3	0.013(0.069)	0.582
Safety Attribute	PUT1	0.035(0.041)	0.393
	PUT2	**0.074(0.033)**	0.023
	PUT3	0.034(0.039)	0.393
Convenience	PUT1	0.011(0.032)	0.744
	PUT2	**0.052(0.025)**	0.034
	PUT3	**0.077(0.029)**	0.008
Comfort	LWuse1	0.168(0.139)	0.227
	LWuse2	**0.226(0.123)**	0.067
	LWuse3	**-0.495(0.110)**	<0.001

Estimates in bold font are statistically significant.

Nevertheless, current usage of the DS did not have consistent impacts on the service quality evaluations as expected, especially regarding the degree of general DS trip comfort. This is probably because the third interview was carried out in winter, and the weather condition (i.e. temperature and daylight hours) was worse than the first two waves. For this reason, in the third wave, people who were actively using the service were more likely to get a lower perception of comfort with the driverless shuttle riding. The explanation for this is supported by a study that highlighted the effects of weather variables on travel satisfaction in Sweden (Ettema et al., 2017). Ettema et al. proposed that winter in Sweden is characterized by cold temperature and few hours of daylight, while sunshine and higher temperatures make commuter more relaxed and satisfied. Therefore, the negative influence from the current usage of the shuttle on comfort perceptions in the third wave can be convincingly explained by the season factor.

However, it is proposed by many studies that the increasing use experience would reinforce users' positive evaluations of the service. Thus, a complement finding was obtained by taking account of cumulative times of ride as TIC in the following section, verifying the expected positive effects of ride experience on the comfort evaluations of DS.

5.1.2.2 Coefficients of Time-invariant Covariates (TICs)

After removing the effects of TVCs on repeated measures, the remaining parts of intercept and slope were regressed on TICs. Except for the TICs mentioned earlier, specific attention needed to be paid to the cumulative times of ridden, which was supposed to develop its value over time. The scale setting of the response options gave rise to an unobvious change in the short term, so only the last measure during the survey period was treated as a time-invariant covariate. This approach of slicing TVC up to a single point in time and treating it as a TIC has been stated (Bollen and Curran, 2006, pp.191).

Table 11 TICs Coefficients Estimates (Standard Errors) for Conditional Models

Model	Predictor Variable	Intercept Factor $\hat{\gamma}_{\alpha m\_n}$	P	Slope Factor $\hat{\gamma}_{\beta m\_n}$	P
Functional Attributes	Intercept	**1.512(0.068)**	0.000	**0.104(0.028)**	0.000
	Gender	-0.046(0.059)	0.435	**0.037(0.020)**	0.063
	Age	0.036(0.064)	0.572	**-0.036(0.022)**	0.095
	Education level	0.003(0.069)	0.970	**-0.057(0.023)**	0.013
Safety Attribute	Intercept	**3.922(0.140)**	0.000	**-0.154(0.065)**	0.018
	Employment status	-0.169(0.109)	0.122	**0.090(0.048)**	0.061
Convenience	Intercept	**3.104(0.098)**	<0.001	**-0.090(0.043)**	0.034
	Gender	-0.106(0.078)	0.172	**0.066(0.030)**	0.028
	Education level	**-0.220(0.090)**	0.015	**0.063(0.035)**	0.076
	Familiarity with autonomous driving technology	**0.164(0.095)**	0.084	-0.021(0.037)	0.575
Comfort (All)	Intercept	**3.661(0.080)**	0.000	**-0.083(0.033)**	0.013
	Car ownership	**-0.157(0.088)**	0.075	0.051(0.037)	0.163
	Cumulative times of ridden	-0.002(0.059)	0.977	**0.064(0.025)**	0.010
Comfort (non-adopters)	Intercept	**3.573(0.051)**	<0.001	**-0.061(0.022)**	0.005
	Familiarity with autonomous driving technology	-0.104(0.125)	0.402	**0.130(0.054)**	0.016
	Attitude towards Technology	**0.086(0.042)**	0.038	-0.020(0.018)	0.271

Estimates in bold font are statistically significant;
n = gender, age, education, familiarity with autonomous driving technology, car ownership, cumulative times of ridden, attitude towards technology;
m = 1 for functional attributes, 2 for safety attribute, 3 for convenience, 4 for comfort and 5 for non-adopters' comfort.

Functional evaluations with age, gender, education level

As shown in Table 11, none of the TICs has significant effects on the intercept factor for functional evaluations, indicating that respondents' initial evaluations of speed and travel time did not vary from individual to individual regarding gender, age, and education level. On the contrary, concerning the explanation of the slope factor, gender, age, and education level all have significant estimates. First, gender significantly explained the slope factor with a positive estimated coefficient. Then, both age and education level have negative and significant estimated coefficients. The findings suggest that males had the same initial ratings as females but eased their ratings faster concerning the functional performance of driverless shuttles. This argument is in close agreement with the previous studies. (Roche-Cerasi, 2019) summarized that males, younger people and high-educated individuals have a more positive attitude towards AV. However, there was no evidence in this study implying the expected effects of neither age nor education level. In contrast, starting with the same ratings, people older than 45 years old changed their attitudes more quickly in the positive direction than younger people. Although this result seems adverse to (Roche-Cerasi, 2019) statement, it aligns with the conclusion that the elderly were more positive toward the characteristics of the driverless shuttle (Nordhoff et al., 2018). Regarding education level, highly educated people rated the speed and travel time of DS in the same way as others at the beginning and raised the scores less, while they rated on-board convenience higher initially and reduced the scores slowly. Therefore, it is safe to suggest that high-educated people keep their ratings on service characteristics more steadily than others.

Safety attribute with employment status

With regard to the safety attribute, employment status was found to have a significant predictive effect on the growth factors. It is revealed that employed respondents reduced their ratings of the safety performance of DS slower than the unemployed group. Hence, we argue that respondents all started with great faith followed by a slow recession in the safety of driverless shuttles, whereas the employed status helped alleviate this recession trend in the latter part of the survey period.

On-board convenience with gender, education, familiarity with autonomous driving technology

Similarly, the influences of gender and education level on service evaluations were also found when respondents assessed on-board convenience. The coefficients suggest that men initially rated on-board convenience of DS the same way as women did, but with a lower rate in the falling tendency of their ratings. This propensity of resistance to decline was also found for the high-educated group compared with others. However, high-educated people were likely to feel slightly unsatisfactory when they take the shuttle for the first time. Nevertheless, this was counteracted if the respondent knew about autonomous driving technology supporting driverless shuttles.

It is noteworthy to mention that after taking account of 'PUT', the effects of these three TICs appeared with a significant and negative mean slope for the remaining observed measures. It turned out that convenience on board and general comfort both had a downward trend in their satisfaction ratings. If we consider these two indicators representative enough, it hints that the overall evaluation of hedonic attributes was declining over the trial period.

General comfort with car ownership and cumulative ridden times

A minor adjustment was made to the conditional model for 'Comfort'. That is, the effects of 'LWuse' (user activity) instead of 'PUT' (public transportation use frequency) was removed from observed measures firstly, and then the rest part was regressed on car ownership and cumulative times of riding.

As we can see in Table 11, it is proved that car owners had a lower initial rating of general comfort compared to respondents who do not own a car in their households. Thus, the results indicate that people who have access to private mobility (i.e. owning at least one car in the household) appear to feel a lower degree of comfort of the shuttle service. This argument is in close agreement with a previous finding that licensed drivers were less willing to utilize shared AVs frequently (Bansal et al., 2016; Nordhoff et al., 2019).

Furthermore, cumulative times of ridden significantly explained the variability in slope factor. This implies that though all respondents reduced their evaluation over the period, the more times they have ridden AB, the slower their evaluation of general comfort decrease over time. Since the earliest inclusion of experience in the UTAUT (Venkatesh et al., 2003), many researchers have validated that increasing exposure to the system leads to a more positive attitude towards it. This is also confirmed in this study by the relation between cumulative times of riding and resistance to the trend of decreasing ratings of comfort. That is to say, adopters who use the service more than five times did not reduce their evaluation of comfort over time, and even raise the scores with more times of ridding.

General comfort of non-adopters and familiarity with autonomous driving technology and attitudes towards technology

All the aforementioned conditional models were modeling on all respondents in general, but non-adopters as a large group of potential users are also of the research interest. To understand the predictive effects of knowledge and attitudes towards technology for the perception of general comfort of non-adopters, we selected four items and incorporated them into a latent construct 'Attitude towards Technology'. All the factor loadings shown in Table 12 exceed 0.60, indicating enough reliability of the measurement model. Importantly, the reliability construct was tested by the CR value of 0.88 (larger than the reasonable value of 0.60), while convergent validity was confirmed by the AVE value of 0.64 (larger than the cutoff value of 0.50). In addition to this latent construct, the familiarity with the underlying technology supporting the driverless shuttles was likewise included.

Table 12 Reliability and Validity Test of Latent Construct 'Attitudes towards Technology'

Indicator	Question Statement	Factor Loading	AVE	CR
W3Q12.2	It's important for me to follow the technology development.	0.73		
W3Q12.3	I'm always among the first to try to use new technology products.	0.84		
W3Q12.4	I'm excited about the possibilities with new technology.	0.86		
W3Q12.5	I know more than others about the latest technology products.	0.80	0.64	0.88

Seven-point Likert scale from 1- completely disagree to 7- completely agree

Being in line with the model for all respondents, trajectories of non-adopters for the general comfort also started with a satisfactory perception and declining over time. However, according to the estimates in Table 11, in line with the well-proven effect of technology savviness highlighted in the MAVA (Nordhoff et al., 2019), individual personality regarding the attitude towards technology and the knowledge about driverless technology has been verified with their positive impacts on the comfort evaluations of the shuttle service. The findings indicate that a good knowledge of autonomous technology results in an increasing trend (instead of decreasing) of the dynamic changes in respondents' evaluations regarding ride comfort. At the same time, a general passion for innovative technology gives rise to a slightly higher initial perception of comfort. On balance, non-adopters with more enthusiasm for technology tend to perceive a higher comfort level of driverless shuttles with a rising trend of their evaluations over time. Therefore, targeted propaganda tactics should be applied to the tech-savvy user groups. For example, operators can cooperate with technology bloggers and Youtubers to create a good atmosphere for the DS service on the internet. When these self-media share their ride experience to the social network platforms, they help to improve the public's awareness of the latest technological upgrades of shared automated vehicles. Subsequently, people who are passionate about emerging technologies will generate interest in having physical using experience with the shuttle service. Ultimately, the future adoption rate among tech-savvy people is expected to be improved.

5.2 Temporal Pattern of User Acceptance

Two longitudinal structural equation models (LSEM) specified for adopters and non-adopters were separately estimated. In this section, hypotheses H1-H6 and H1'-H5' in the user acceptance framework portrayed in Section 4.2.1 (Figure 10 and Figure 11) will be tested. Besides, moderator variables were also taken into account for the non-adopters model to explore what factors can encourage the adoption intention of potential users.

Being supported by the analysis of the latent growth curve for all respondents, three service quality attributes were treated as independent variables in LSEM models. Along with the other three categories of variables (i.e. dependent, mediator and moderator variables), the means of the indicators in two waves for the two groups, adopters and non-adopters, are displayed in Table 13, respectively. Two time points refer to the second and third survey wave, corresponding to initial propensity and continuance propensity, respectively.

The development patterns of independent variables were revealed by the latent growth curve analysis. Nonetheless, the differences in means of BI and WOM did not significantly depart from zero for both groups (due to the relatively short survey period). Hence, the changes in dependent variables can only be understood by following statistics. For adopters, 50 people (51%) maintained their continuance intention to use the driverless shuttle while 22 people (22%) enhanced meagerly. For non-adopters, 115 people (52%) remained, and 54 people (24%) increased their ratings on the adoption intention. Likewise, the willingness of word-of-mouth recommendation kept the same as before for 54% of adopters and 23% of non-adopters, while the willingness increased for 24% of adopters and 20% of non-adopters.

Table 13 Means Statistics of Independent, Dependent, Mediator and Moderator Variables

		Adopters		Non-adopters	
Indicator	Statement	Wave2 Mean	Wave3 Mean	Wave2 Mean	Wave3 Mean
Independent Variable					
FA_item1	Speed of the driverless shuttle compared to a regular bus	1.612	1.745	1.589	1.770
FA_item2	Travel time of driverless shuttle compared to a regular bus	1.898	2.041	1.992	2.068
FA_item3	Travel time of driverless shuttle compared to a car	1.755	1.847	1.725	1.815
SA_item1	Perceived safety when there is a driver/steward on the DS	3.786	3.714	3.804	3.581
HA_item1	General comfort of driverless shuttle trip	3.684	3.459	3.509	3.434
HA_item2	Convenience on the driverless shuttle	3.020	2.898	3.102	3.174
HA_item3	Comfort on board due to driving speed and driving pattern of the driverless shuttle	3.398	3.194	3.438	3.340
Dependent Variable					
BI	Would you use self-driving buses more often in the future?	3.622	3.551	3.623	3.626
WOM	Will you recommend DS to your colleagues/friends/family?	3.561	3.612	3.336	3.294
Mediator Variable					
Satisfaction	Satisfaction of the driverless shuttle trip	3.673	3.633		
PU_item1	Prospect of the importance of the future role of the driverless shuttle in the transport system			3.928	3.917
PU_item2	Driverless shuttle would be suitable for daily travel.			4.011	3.864
PU_item3	Driverless shuttle would meet respondents' travel needs.			3.264	3.200
Moderator Variable					
PUT	Weekly use frequency of public transport connecting to the transit hub			2.525	2.777
Tech_item1	It's important for me to follow the technology development.				4.815
Tech_item2	I'm always among the first to try to use new technology products.				3.547
Tech_item3	I'm excited about the possibilities with new technology.				3.826
Tech_item4	I know more than others about the latest technology products.				3.423

All the items are in the five-point Likert scale. Only Tech_item1 to Techh_item4 are in the seven-point Likert scale.

5.2.1. Measurement Model Results

Table 14 Results of Reliability and Convergent Validity Examination

WⅡ Construct	Indicator	Factor Loading	AVE	CR	WⅢ Construct	Indicator	Factor Loading	AVE	CR
Functional Attribute2	W2Q27	0.581			Functional Attribute3	W3Q36	0.632		
	W2Q36	0.798				W3Q44	0.845		
	W2Q37	0.825	0.55	0.78		W3Q45	0.766	0.57	0.79
Hedonic Attribute2	W2Q30	0.644			Hedonic Attribute3	W3Q38	0.632		
	W2Q31	0.683				W3Q39	0.739		
	W2Q32	0.633	0.43	0.69		W3Q40	0.66	0.46	0.72
Perceived Usefulness2	W2Q17	0.735			Perceived Usefulness3	W3Q28	0.745		
	W2Q22	0.845				W3Q31	0.865		
	W2Q23	0.641	0.56	0.79		W3Q32	0.717	0.61	0.82

Table 15 Results of Discriminant Validity Examination

WⅡ_sqrt(AVE)/Correlation	FunctionalAttr2	HedonicAttr2	Usefulness2
FunctionalAttr2	0.74		
HedonicAttr2	0.289	0.66	
Usefulness2	0.347	0.531	0.75
WⅢ_ sqrt(AVE)/Correlation	FunctionalAttr3	HedonicAttr3	Usefulness3
FunctionalAttr3	0.75		
HedonicAttr3	0.269	0.69	
Usefulness3	0.267	0.42	0.78

Three latent variables were included in the adopters' and non-adopters' model at each wave. In addition to the mediating variable 'Usefulness', two explanatory variables, 'Functional Attributes' and 'Hedonic Attributes' were also measured by three indicators, respectively. Consequently, a total of six measurement constructs at two time points were examined in the CFA section.

Results listed in Table 14 indicate that all the factor loadings are beyond 0.60 except for the indicator 'W2Q27', which has an AVE value of 0.58. Nevertheless, 0.58 is considered extremely close to 0.6 thus is acceptable. Therefore, the reliability of each indicator is confirmed. Furthermore, all values of CR exceed the criterion of 0.60, verifying the internal consistency of the measurement model. AVE values for 'Functional Attributes' and 'Perceived Usefulness' are all larger than the cutoff value of 0.50, proving the convergent validity. In respect to the 'Hedonic Attribute', although AVEs do not reach the cutoff value, 0.43 and 0.46 are considered acceptable in the context of current empirical research. Lastly, discriminant validity was examined by comparing the inter-construct correlation and the square root of AVEs. As displayed in Table 15, all the values in the lower-left corner (inter-construct correlations) are less than the ones on the diagonal (square root of AVEs), confirming the discriminant validity of the model. Additionally, multicollinearity and multivariate normality were tested in Section 4.2.2. Hence, the basic hypotheses of ML estimation method were verified.

5.2.2. Structural Regression Model Results

Model fit of the LSEM models for adopters and non-adopters was examined. Table 16 lists the summary of fit indices of the hypothesis models. Both models have χ^2_M/df ratio less than 5 while CFI and TLI values are larger than 0.90, which providing evidence of acceptable fitting models. Next, GFI values are both larger than 0.80, showing an acceptable fit as well. The value of RMSEA for non-adopters' model equals to 0.05, implying a satisfactory fit. Although the RMSEA value of 0.07 for the adopters' model is higher than the recommended threshold of 0.05, it still indicates a reasonable error of approximation. According to the threshold values reviewed in Section 4.2.2, it is reasonable to argue that both models achieved satisfactory or acceptable fit.

In two longitudinal models, the explained variances in dependent variables BI and WOM are shown in Table 17. Due to the inclusion of sequential updating effects of propensities at the previous time point, higher variances were explained in the continuance using intention (BI3) and continuance willingness to recommend to others (WOM3) for both adopters' model and non-adopters' model.

Table 16 Model Fit Indices of the Hypothesis Models

	Chi-square Ratio (χ^2_M/df)	Goodness of Fit Index (GFI)	Comparative Fit Index (CFI)	Tucker-Lewis Index (TLI)	Root Mean Square Error of Approximation (RMSEA)
Adopters	1.47	0.82	0.92	0.9	0.07
Non-adopters	1.66	0.87	0.93	0.91	0.05

Table 17 Explained Variance in Behavior Intention and Word-of-mouth Recommendation

	Adopters' Model		Non-adopters' Model	
	BI	WOM	BI	WOM
Wave2	0.30	0.55	0.56	0.47
Wave3	0.50	0.65	0.61	0.49

For adopters

Figure 14 displays the estimated path diagram modeling the longitudinal changes in adopters' acceptance of the driverless shuttle. Based on the analysis of the dynamic changes in service evaluations in the LGCM section, three service attributes were supposed to impact overall satisfaction. The results indicated that functional attributes with low performance did not significantly influence satisfaction at either of the time points. Hence, hypothesis 1 was rejected at both initial time point and subsequent time point. This finding suggested that the lower speed and longer travel time of DS compared with regular bus and car do not necessarily reduce user satisfaction with the service. In other words, the public has a relatively high tolerance to a slow driverless shuttle trip in the trial operation phase. At the same time, the safety attribute only played a decisive role in the initial valuation of the overall satisfaction but made no significant contribution to the current satisfaction. Therefore, hypothesis 2 was verified at the initial time point, whereas it was not accepted at the subsequent time point. Accordingly, the results implied that when people feel safe on the driverless shuttle, they are more likely to assess the trip as satisfactory initially. However, as time goes by, people's safety concern becomes less critical for the subsequent user satisfaction. On the other hand, hedonic attributes most significantly predicted the adopters' overall satisfaction with the driverless bus ride experience in both stages, implying that hypothesis 3 was consistently true for adopters. This means that perceptions of hedonic attributes are the most significant determinants of user satisfaction, and thus a more comfortable and convenient prior and current travel experience could make users more satisfied with the overall service quality.

Next, both initial using intention (BI2) and continued using intention (BI3) were significantly affected by the degree of overall satisfaction, which verified hypothesis 4. Therefore, it is reasonable to submit that satisfaction indeed has mediating effects on adoption intention. The finding implies that when individuals have a satisfactory assessment of the ride experience, they are expected to adopt the DS service more frequently in the future. Additionally, since satisfaction is supposed to be a mediator, perceptions of comfort, convenience and safety concerns that determine satisfaction were considered indirect determinants of adoption intention. Hence, it is reasonable to suggest that a more comfortable, convenient and safe ride experience could help retain users' adoption. Similarly, satisfaction is demonstrated as a direct determinant of the willingness to recommend to others, and this causal effect was also found at both time points (H5 was accepted). The results revealed that users with positive valuations of the service are more likely to spread their appreciation with the DS to their friends, family members or colleagues. In addition, using intention had a significant positive effect on the word-of-mouth recommendation activity, so that hypothesis 6 was verified among the adopters. This finding indicated that users who intend to use the DS service in the future are also willing to recommend the service to others, which in turn may give rise to an improvement of adoption rate among potential user groups.

Meanwhile, our results indicate that all variables were significantly influenced by the perception or evaluation from the previous time point. These intertemporal relationships can be explained by the sequential updating mechanisms (Kim and Malhotra, 2005). Kim and Malhotra shed light on the sequential updating mechanism underlying postadoption phenomena based on the TAM framework in the longitudinal technology acceptance context. In line with the finding of Kim and Malhotra, the sequential updating mechanism was also validated in the current study. Our results implied that when people give the newly formed evaluations of DS service, they refer to prior evaluations as an anchor and regard new perceptions of the service quality as adjustment. Notably, through the intertemporal influencing paths, the earlier satisfaction indirectly but robustly affected both continued adoption intention and continued willingness of WOM recommendation. This phenomenon further confirms that the public's first impression of the shuttle service is prominent to encourage continued adoption in the long term.

For non-adopters

Similar to the way representing results of the adopters' model, Figure 15 shows the resulting path diagram of non-adopters' model. Due to the lack of actual ride experience, satisfaction was replaced by perceived usefulness as a substitute for non-adopters. On top of the explanatory variables relating to service quality, moderators were included in the model to help identify potential user groups and develop strategies and corresponding tactics to encourage future adoption.

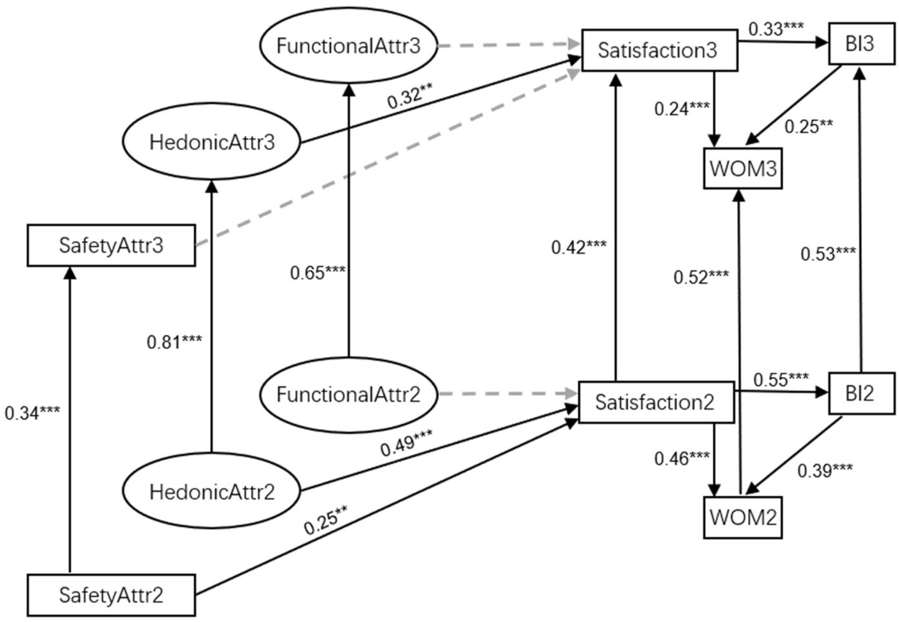
The moderators were selected from the influencing factors in the LGCM conditional model. Among them, 'PUT' and 'attitude towards technology' showed significant results and were involved in the final model.

Different from the adopters' model, the three service attributes all presented significant decisive influences on the perception of usefulness, indicating that H1', H2' and H3' could be verified at the initial time point. Accordingly, the results reveal that the initial perception of usefulness depends on the individual's comprehensive assessment of driverless shuttles' hedonic, functional and safety performance. In other words, only when the degree of comfort, convenience, sense of safety and speed of the driverless shuttle all reach people's requirements, they expect the driverless shuttle to meet their daily travel needs and admit its essential role in the transport system. However, the effects of functional attributes and safety attribute became insignificant afterwards, indicating that H1 'and H2' were not accepted at the sequential time point. Thus, our results suggest that without actual ride experience, the updates of perceptions regarding the speed and safety of the driverless shuttle became less important than individual's earlier assessment for them to evaluate the expected usefulness of the service again. With this in mind, it has become increasingly important to provide a more comfortable, safer, and faster trip, enhancing the appealingness of driverless shuttles to prospective users. Notably, only hedonic attributes were again verified as consistently significant, supporting to accept H3' at both time points. This finding embraces the argument that a comfortable and convenient trip is the most motivating factor for non-adopters to appreciate the DS service as well.

Subsequently, positive effects of perceived usefulness on BI and WOM were both confirmed, leading to the acceptance of H4' and H5'. This finding revealed that people tend to feel it more reasonable to adopt the service and spread this impression to their social networks when they agree that the DS service can meet their travel needs. Nevertheless, because of the insignificant relationship between BI and WOM, H6 was failed to be accepted in the non-adopters group. The results implied that people who are more willing to adopt the driverless shuttle might not recommend the service to others simultaneously. It is possible to argue that the actual ride experience of influencer is an essential prerequisite for the word-of-mouth marketing of the driverless shuttle service. Therefore, to create a favorable public opinion atmosphere and increase the market share of the service, it is imperative to improve the adoption rate at the early stage of operation.

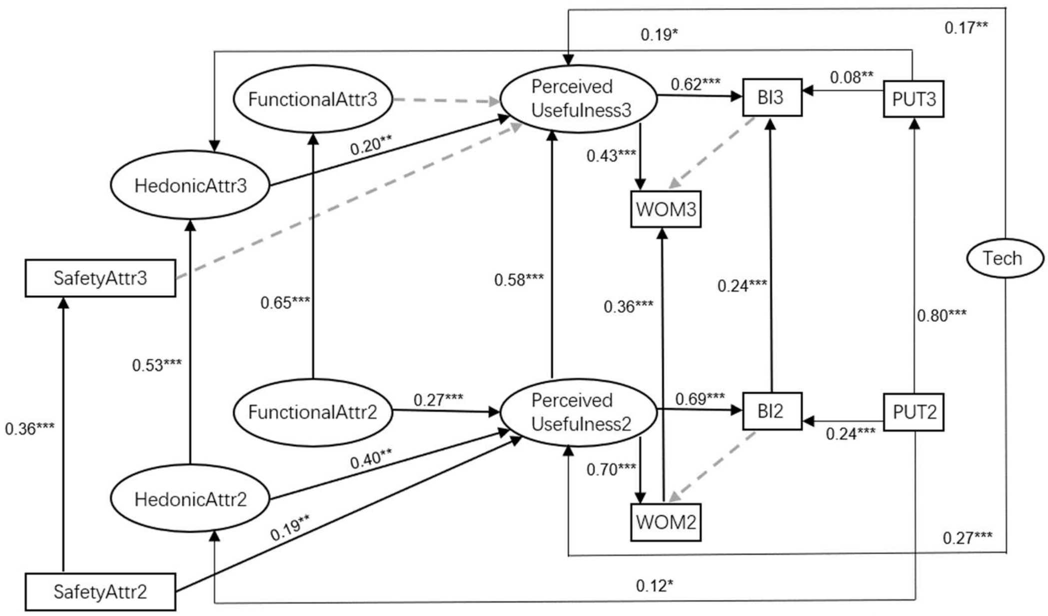
In addition, the coefficients of intertemporal paths of service quality evaluations, perceived usefulness, BI and WOM from earlier to later time point are all significant. Therefore, the sequential updating effects are demonstrated among non-adopters as well. Subsequently, indirect effects of earlier perceived usefulness on BI and WOM at the sequential time point again proved the significant role of the initial evaluation in long-term public acceptance. The results suggest that although people may decide to use driverless shuttles in the future and fail to act on it, their initial perceptions and evaluations of the service will continue to shape their future intentions in the long run.

In particular, the use of public transport as a feeder and individual attitude toward technology are also included in the model. The frequency of public transport usage ('PUT') varies from time to time. Nevertheless, 'PUT' directly and positively contributed to the propensity of adopting and recommending the shuttle service through its direct effects on the perception of the hedonic attribute in each wave. As a consequence, public transport passengers were identified as potential groups who are more receptive to driverless shuttles. Next, the indirect effects of individual attitude towards technology on perceived usefulness were also projected on BI and WOM through the mediating effect. This indicates that people who have passions for innovative technology tend to perceive the shuttles as more practical, and in turn, more likely to adopt the service in the future. In general, frequent users of public transportation and tech-savvy people are potential groups who are more likely to accept driverless shuttles. After identifying these users, targeted campaigns and tactics should be tailored.



Significant estimate with * p-value< 0.10; ** p-value<0.05; *** p-value <0.001.
Insignificant estimate with the dotted line.

Figure 14 Changes in Adopters' Acceptance of the Driverless Shuttle



Significant estimate with * p-value< 0.10; ** p-value<0.05; *** p-value <0.001.
Insignificant estimate with dotted line.

Figure 15 Changes in Non-adopters' Acceptance of the Driverless Shuttle

6. Discussion and Conclusion

To successfully switch driverless shuttles from the demonstration stage to the feeder service on a regular basis, it is imperative to continuously monitor the public's attitudes affecting user acceptance. However, most of the previous studies were based on cross-sectional data, lacking an understanding of the temporal changes in the long term. Therefore, this thesis aims at filling the gap in the longitudinal analysis of people's intention to adopt a regular-scheduled driverless shuttle service operated in a mixed traffic environment on public roads.

This study was designed to examine the change in user attitudes toward a regularly scheduled driverless shuttle service over time and identify the determinants of intention to adopt and recommend the service. Specifically, to understand the continued intention to use the shuttles and continued willingness of word-of-mouth recommendation among adopters and non-adopters, we embedded the hierarchy of service quality attributes adapted from the three-factors theory into the user acceptance model by utilizing two-wave panel data extracted from a three-wave panel survey. To our knowledge, this study is the first attempt of exploring the effects of service quality on behavior intention and willingness of recommendation through the mediating effects of service satisfaction (perceived service usefulness) in a driverless shuttle context over time.

Through merging the valuations of service attributes into the user acceptance model, two longitudinal models were developed for adopters and non-adopters separately. Satisfaction of adopters and perceived usefulness of non-adopters were found to positively affect people's adoption intention as well as their favorable word-of-mouth behavior as hypothesized. Especially when looking at the intertemporal relationships, prior satisfaction (perceived usefulness) was also confirmed as essential for the continuance intention. Furthermore, by viewing the driverless shuttle as a public transport mode, it is proved that the hierarchical evaluation of the transit service quality contributed to adopters' satisfaction and non-adopters' perception of usefulness, but the judging criteria of evaluations changed over time. Therefore, in line with the statement about the substantial effect of service characteristics on user acceptance of driverless shuttle in a post-riding interview study conducted by Nordhoff, Winter, et al. (2019), the current study verified that service quality evaluation is an essential antecedent of public acceptance of the driverless shuttle service.

Our findings suggest that the overall satisfaction with driverless shuttle's service quality consist of hedonic, safety and functional three service attributes. In the current study, respondents evaluated functional attributes relatively low and increased steadily, while they perceived safety and hedonic attributes as satisfactory, but the assessments declined moderately over time. Consequently, to encourage the future adoption of driverless shuttles, it is vital to promote favorable evaluations in the suitable aspect of service quality.

No matter for adopters or non-adopters, improving the performance of comfort and convenience on-board of driverless shuttles is the most critical measure to improve the future adoption rate. For those who have actually taken the driverless shuttle, a more comfortable and convenient prior and current travel experience could make them more satisfied with the overall service quality and become the primary motivation to continue adopting the service. They will even be willing to convey this good impression to their family and friends, so as to create a publicity atmosphere that providing a higher degree of exposure to the shuttle service for more potential users. On the other hand, from the perspective of non-adopters, continuous pleasant perceptions of hedonic attributes can lead to a higher expected utility of the shuttles. Thereupon, people tend to feel it more reasonable to adopt the service when they reach an agreement that the driverless shuttle service can meet their travel needs while playing an important role in the existing transportation system. At the same time, more attention should be paid to the perceptions of the hedonic attributes by those who do not frequently use public transport to connect to the transit hub. Enhancing car owners' initial impressions of the shuttle, for example, could encourage them to shift from driving a private car to a comparatively more comfortable and convenient driverless shuttle riding.

The second decisive factor contributing to the valuation of satisfaction is the initial trust in the safety of the driverless shuttle. We believe that the auxiliary presentations of emergency plans for specific incidents will enhance the sense of security as the first impression for the new adopters. In addition, a steady transition from backing up by a steward to completely no human intervention on board will help to maintain the present perceived safety of passengers.

Lastly, the comparison of the speed and travel time of driverless shuttle with other transport modes is more important for non-adopters. In accordance with previous findings (Roche-Cerasi, 2019), on average, perceived functional performance is relatively low when comparing the shuttles to existing transport modes. According to the three-factor theory, respondents' dissatisfaction should negatively affect the overall evaluation of the service. However, a positive relation was found between functional attribute and non-adopters' initial perception of usefulness, implying the public's high tolerance to a relatively slow driverless shuttle trip in the trial operation phase. In order to take this advantage, the policies should support the shuttles to operate at higher speeds for a shorter travel time in response to public expectations as soon as possible.

In general, current public transport users and tech-savvy people are the potential users who are more likely to adopt the shuttle service. After identifying these user groups, targeted measures can be developed to better stimulate the conversion of potential users to adopters. For instance, a special discount on SL ticket for the driverless shuttle ride can be appealing to public transport users, leading to the shift from regular bus to driverless shuttles. Or, decision-makers can collaborate the driverless shuttle with other available transportation modes in the monthly package of MaaS, providing seamless and integrated intermodal travel. Alternatively, we also suggest displaying more updating achievements of the shuttles in e-journals and cooperate with self-media (e.g. technology bloggers and Youtubers) to share their trial ride experience to the internet, so that more people who are passionate about emerging technologies will have a better awareness of the latest technological upgrades of shared automated vehicles, and thus, generate interest in having physical using experience with the shuttle service. Additionally, the moderator analysis implies that male, aged beyond 45, high-educated and employed are the user profile's specific characteristics that bring positive influence onto the rate of changes in respondents' ratings of service quality. Meanwhile, car owners and high-educated groups were less satisfied when they were interviewed about the shuttle's service quality for the first time. Lastly, the higher degree of interaction with the driverless shuttle, the more favorable evaluation people would give.

Limitations and future research directions

There are several limitations in the current study. Firstly, the actual behavior was not included in the user acceptance model due to the short trial period and the setting of response scales in the questionnaire. Instead, behavioral intention was treated as the dependent variable reflecting how acceptant the public perceives the driverless shuttle service. Therefore, next step, we should take into account the changes in actual usage of shuttles by setting a more specific response scale for the panel survey with a longer time interval between waves in the future.

Second, the ideal approach to understand the longitudinal pattern of user acceptance is supposed to form a multivariate latent growth curve model (MLCM) for the regression analysis. Nevertheless, the questions relating to BI and WOM were only interviewed in the last two waves. Alternatively, this study applied longitudinal structural equation models to the two-wave panel data with a compensating LGCM analysis of explanatory variables based on the complete three-wave data. Thus, MLCM is considered a future advancement to provide a deeper understanding of the individual variation of driverless shuttle's public acceptance over time in a way that minimizes the loss of information.

Third, the hypothesized relationships between variables were examined with restricted numbers of observed items (i.e. safety attribute was measured by single item) and based on a relatively small-sized (363 responses) sample, which might lead to implications not as expected. In future research, it should be important to refine and tailor the items measuring service quality and attempt to examine the hypotheses within a larger and more representative sample.

Although the conclusions in the current study are drawn with limitations, it is expected that this framework integrating the service quality attributes, customer satisfaction and user acceptance theory together and could be helpful to develop tactics aiming at diffusion and generalization of the driverless shuttle service.

Appendix A

Assessment of normality (Non-adopters)						
Variable	min	max	skew	c.r.	kurtosis	c.r.
W2Q7	1	4	-0.082	-0.548	-1.234	-4.101
W2Q28	1	5	-0.218	-1.451	-0.349	-1.159
W3Q25	1	4	-0.277	-1.842	-1.322	-4.392
W3Q37	1	5	-0.29	-1.928	0.002	0.007
W2Q24	1	5	-0.737	-4.899	0.213	0.709
W2Q19	1	5	-0.355	-2.357	-0.092	-0.306
W3Q12.3	1	7	0.147	0.976	-0.722	-2.398
W3Q12.2	1	7	-0.563	-3.738	-0.582	-1.933
W3Q12.5	1	7	0.113	0.75	-1.162	-3.862
W3Q12.4	1	7	0.019	0.128	-1.038	-3.448
W3Q29	1	5	-0.312	-2.075	0.487	1.619
W2Q17	1	5	-0.663	-4.405	-0.115	-0.382
W3Q28	1	5	-0.564	-3.746	0.189	0.627
W2Q32	1	5	-0.462	-3.07	0.542	1.8
W3Q40	1	5	-0.171	-1.138	0.162	0.539
W2Q22	1	5	-0.587	-3.903	-0.405	-1.345
W2Q23	1	5	-0.267	-1.774	-0.762	-2.531
W3Q31	1	5	-0.545	-3.625	-0.024	-0.079
W3Q32	1	5	-0.155	-1.028	-0.398	-1.322
W3Q33	1	5	-0.77	-5.117	0.543	1.805
W3Q45	1	5	1.085	7.212	0.695	2.309
W2Q37	1	5	1.198	7.959	1.032	3.428
W2Q27	1	4	1.205	8.007	0.925	3.073
W2Q36	1	5	0.869	5.774	0.7	2.327
W3Q36	1	4	0.618	4.11	-0.375	-1.245
W3Q44	1	5	0.689	4.576	0.05	0.166
W2Q30	1	5	-0.125	-0.832	0.084	0.278
W2Q31	2	5	0.546	3.631	1.215	4.037
W3Q39	1	5	-0.011	-0.074	1.278	4.247
W3Q38	1	5	-0.573	-3.808	1.146	3.807
Multivariate					88.596	16.457

Assessment of normality (adopters)						
Variable	min	max	skew	c.r.	kurtosis	c.r.
W2Q28	1	5	-0.66	-2.669	0.251	0.507
W2Q5	1	5	-0.666	-2.69	0.243	0.491
W3Q37	1	5	-0.502	-2.031	-0.178	-0.359
W3Q23	1	5	-0.668	-2.702	0.182	0.369
W2Q24	1	5	-0.739	-2.986	0.136	0.275
W2Q19	1	5	-0.642	-2.594	-0.288	-0.583
W3Q33	1	5	-0.77	-3.114	-0.18	-0.364
W3Q29	1	5	-0.563	-2.274	-0.148	-0.3
W2Q32	1	5	-0.576	-2.328	0.038	0.076
W3Q40	1	5	-0.315	-1.272	-0.25	-0.504
W3Q45	1	5	1.171	4.731	0.561	1.134
W2Q37	1	5	1.223	4.942	0.862	1.742
W2Q27	1	5	1.545	6.244	1.758	3.552
W2Q36	1	4	0.931	3.762	0.351	0.709
W3Q36	1	5	1.158	4.68	1.581	3.195
W3Q44	1	5	0.931	3.764	0.421	0.851
W2Q30	1	5	-1.126	-4.551	1.617	3.268
W2Q31	1	5	-0.292	-1.178	-0.03	-0.06
W3Q39	1	5	-0.221	-0.892	0.223	0.45
W3Q38	1	5	-0.822	-3.323	-0.021	-0.043
Multivariate					38.152	6.366

Appendix B

Multicollinearity Tests				
	(Adopters)		(Non-adopters)	
	Tolerance	VIF	Tolerance	VIF
W2Q36				
W3Q44	0.534	1.872	0.428	2.336
W2Q37	0.563	1.776	0.672	1.489
W3Q45	0.473	2.116	0.483	2.072
W2Q28	0.852	1.174	0.821	1.218
W3Q37	0.624	1.602	0.779	1.283
W2Q27	0.541	1.848	0.6	1.667
W3Q36	0.453	2.206	0.739	1.354
W2Q31	0.377	2.656	0.682	1.467
W3Q39	0.5	2.002	0.751	1.331
W2Q30	0.541	1.848	0.758	1.319
W3Q38	0.764	1.309	0.815	1.227

References

Abenoza, R. F., Cats, O., & Susilo, Y. O. (2019). Determinants of traveler satisfaction: Evidence for non-linear and asymmetric effects. Transportation Research Part F: Traffic Psychology and Behaviour, 66, 339–356. https://doi.org/10.1016/j.trf.2019.09.009

Adler, R. (2019, December 13). Autonomous or merely highly automated – what is actually the difference? Fraunhofer IESE. https://www.iese.fraunhofer.de/blog/autonomous-or-merely-highly-automated-what-is-actually-the-difference/

Ainsalu, J., Arffman, V., Bellone, M., Ellner, M., Haapamäki, T., Haavisto, N., Josefson, E., Ismailogullari, A., Lee, B., Madland, O., Madžulis, R., Müür, J., Mäkinen, S., Nousiainen, V., Pilli-Sihvola, E., Rutanen, E., Sahala, S., Schønfeldt, B., Smolnicki, P. M., . . . Åman, M. (2018). State of the Art of Automated Buses. Sustainability, 10(9), 3118. https://doi.org/10.3390/su1009-3118

Alexander-Kearns, M., Peterson, M., & Cassady, A. (2016). The impact of vehicle automation on carbon emissions. Center for American Progress. Retrieved from Https://www. Americanprogress.Org/issues/green/reports/2016/11/18/292588/theimpact-of-Vehicle-Automation-on-Carbon-Emissions-Where-Uncertainty-Lies.

Allen, J., Muñoz, J. C., & Ortúzar, J. d. D. (2019). Understanding public transport satisfaction: Using Maslow's hierarchy of (transit) needs. Transport Policy, 81, 75–94. https://doi.org/10.1016/j.tranpol.2019.06.005

Awang, Z. (repr. 2013, 2012). Structural equation modeling using amos graphic. UiTM Press.

Azad, M., Hoseinzadeh, N., Brakewood, C., Cherry, C. R., & Han, L. D. (2019). Fully Autonomous Buses: A Literature Review and Future Research Directions. Journal of Advanced Trans-portation, 2019, 1–16. https://doi.org/10.1155/2019/4603548

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. Journal of the Academy of Marketing Science, 16(1), 74–94. https://doi.org/10.1007/BF02723327

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. Transportation Research Part C: Emerging Technologies, 67, 1–14. https://doi.org/10.1016/j.trc.2016.01.019

Barkarby. (2020). Välkommen till Barkarby | Barkarby. Järfälla municipality. https://barkarby.se/

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in market-ing and consumer research: A review. International Journal of Research in Marketing, 13(2), 139–161. https://doi.org/10.1016/0167-8116(95)00038-0

Becker, F., & Axhausen, K. W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. Transportation, 44(6), 1293–1306. https://doi.org/10.1007/s11116-017- 9808-9

Berg, I., Rakoff, H., Shaw, J., Smith, S. B., & John A. Volpe National Transportation Systems Center. (2020, July 31). System Dynamics Perspective for Automated Vehicle Impact Assessment (FHWA-JPO-20-809). United States. Department of Transportation. Intelligent Transportation Systems Joint Program Office. https://rosap.ntl.bts.gov/view/dot/49813

Bollen, K. A. (1989). Structural equations with latent variables. Wiley series in probability and mathematical statistics. Wiley.

Bollen, K. A., & Curran, P. J. (2006). Latent curve models: A structural equation perspective/ Kenneth A. Bollen, Patrick J. Curran. Wiley series in probability and statistics. Wiley; Chichester : John Wiley [distributor].

Bone, P. F. (1992). Determinants of Word-Of-Mouth Communications During Product Consumption. ACR North American Advances, NA-19. https://www.acrwebsite.org/volumes/7359/volumes/v19/NA-19

Brown, J. J., & Reingen, P. H. (1987). Social Ties and Word-of-Mouth Referral Behavior. Journal of Consumer Research, 14(3), 350–362. http://www.jstor.org/stable/2489496

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. Sociological Methods & Research, 21(2), 230–258. https://doi.org/10.1177/0049124192021002005

Brunner, M., & SÜβ, H.-M. (2005). Analyzing the Reliability of Multidimensional Measures: An Example from Intelligence Research. Educational and Psychological Measurement, 65(2), 227–240. https://doi.org/10.1177/0013164404268669

Chee, P. N. E., Susilo, Y. O., & Wong, Y. D. (2020). Determinants of intention-to-use first-/last-mile automated bus service. Transportation Research Part a Policy and Practice, 139, 350–375. https://doi.org/10.1016/j.tra.2020.06.001

Chee, P. N. E., Susilo, Y. O., & Wong, Y. D. (2021). Longitudinal interactions between experienced users' service valuations and willingness-to-use a first-/last-mile automated bus service. Travel Behaviour and Society, 22, 252–261. https://doi.org/10.1016/j.tbs.2020.10.004

Cheng, S.-I. (2011, June). Comparisons of competing models between attitudinal loyalty and behavioral loyalty. International Journal of Business and Social Science, Vol.2 No.10. http://www.ijbssnet.com/journals/vol._2_no._14%3b_july_2011/18.pdf

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. Journal of Cognition and Development : Official Journal of the Cognitive Development Society, 11(2), 121–136. https://doi.org/10.1080/15248371003699969

David Levinson. (2017). On the Differences between Autonomous, Automated, Self-Driving, and Driverless Cars. Transportist. https://transportist.org/2017/06/29/on-the-differences-between-autonomous-automated-self-driving-and-driverless-cars/

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319. https://doi.org/10.2307/249008

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A Confirmatory Factor Analysis of the End-User Computing Satisfaction Instrument. MIS Quarterly, 18(4), 453. https://doi.org/10.2307/249524

EasyMile. EZ10 passenger shuttle | EasyMile. EasyMile. https://easymile.com/vehicle-solutions/ez10-passenger-shuttle

Ettema, D., Friman, M., Olsson, L. E., & Gärling, T. (2017). Season and Weather Effects on Travel-Related Mood and Travel Satisfaction. Frontiers in Psychology, 8, 140. https://doi.org/ 10.3389/fpsyg.2017.00140

F. Hair Jr, J., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM). European Business Review, 26(2), 106–121. https://doi.org/10.1108/EBR-10-2013-0128

Friedrich, B. (2016). The Effect of Autonomous Vehicles on Traffic. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), Autonomous driving: Technical, legal and social aspects/Markus Maurer, J. Christian Gerdes, Barbara Lenz, Hermann Winner, editors (pp. 317–334). Springer OPen. https://doi.org/10.1007/978-3-662-48847-8_16

Fu, X., & Juan, Z. (2017). Understanding public transit use behavior: Integration of the theory of planned behavior and the customer satisfaction theory. Transportation, 44(5), 1021–1042. https://doi.org/10.1007/s11116-016-9692-8

Harrison-Walker, L. J. (2001). The Measurement of Word-of-Mouth Communication and an Investigation of Service Quality and Customer Commitment As Potential Antecedents. Journal of Service Research, 4(1), 60–75. https://doi.org/10.1177/109467050141006

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multi-disciplinary Journal, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

KANO, N., SERAKU, N., TAKAHASHI, F., & ichi TSUJI, S. (1984, April). Attractive Quality and Must-Be Quality. Journal of the Japanese Society for Quality Control, 14(2), 147–156. http://ci.nii.ac.jp/naid/110003158895/en/

Kellett, J., Barreto, R., van Hengel, A. den, & Vogiatzis, N. (2019). How Might Autonomous Vehicles Impact the City? The Case of Commuting to Central Adelaide. Urban Policy and Research, 37(4), 442–457. https://doi.org/10.1080/08111146.2019.1674646

Kim, S. S., & Malhotra, N. K. (2005). A Longitudinal Model of Continued IS Use: An Integrative View of Four Mechanisms Underlying Postadoption Phenomena. Management Science, 51(5), 741–755. https://doi.org/10.1287/mnsc.1040.0326

Kline, R. B. (2016). Principles and practice of structural equation modeling (Fourth edition). Methodology in the social sciences. The Guilford Press.

Krueger, R., Rashidi, T. H., & Rose, J. M. (2016). Preferences for shared autonomous vehicles. Transportation Research Part C: Emerging Technologies, 69, 343–355. https://doi.org/10.1016/j.trc.2016.06.015

KTH. (2020, December 2). MMiB Modern Mobility in Barkaby | KTH. https://www.itrl.kth.se/research/ongoingprojects/mmib-modern-mobility-in-barkaby-1.917931

Lawley, D. N. (1940). VI.—The Estimation of Factor Loadings by the Method of Maximum Likelihood. Proceedings of the Royal Society of Edinburgh, 60(1), 64–82. https://doi.org/10.1017 /S037016460002006X

Madigan, R., Louw, T., Dziennus, M., Graindorge, T., Ortega, E., Graindorge, M., & Merat, N. (2016). Acceptance of Automated Road Transport Systems (ARTS): An Adaptation of the UTAUT Model. Transportation Research Procedia, 14, 2217–2226. https://doi.org/10.1016/j.trpro.2016.05.237

Mahmoodi Nesheli, M., Li, L., Palm, M., & Shalaby, A. (2021). Driverless shuttle pilots: Lessons for automated transit technology deployment. Case Studies on Transport Policy. Advance online publication. https://doi.org/10.1016/j.cstp.2021.03.010

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. Psychological Bulletin, 97(3), 562–582. https://doi.org/10.1037/0033-2909.97.3.562

Netemeyer, R., Bearden, W., & Sharma, S. (2003). Scaling Procedures. Sage (Atlanta, Ga.). Advance online publication. https://doi.org/10.4135/9781412985772

Nieuwenhuijsen, J., Correia, Gonçalo Homem de Almeida, Milakis, D., van Arem, B., & van Daalen, E. (2018). Towards a quantitative method to analyze the long-term innovation diffusion of automated vehicles technology using system dynamics. Transportation Research Part C: Emerging Technologies, 86, 300–327. https://doi.org/10.1016/j.trc.2017.11.016

Nordhoff, S., Kyriakidis, M., van Arem, B., & Happee, R. (2019). A multi-level model on automated vehicle acceptance (MAVA): a review-based study. Theoretical Issues in Ergonomics Science, 20(6), 682–710. https://doi.org/10.1080/1463922X.2019.1621406

Nordhoff, S., van Arem, B., & Happee, R. (2016). Conceptual Model to Explain, Predict, and Improve User Acceptance of Driverless Podlike Vehicles. Transportation Research Record: Journal of the Transportation Research Board, 2602(1), 60–67. https://doi.org/10.3141/2602-08

Nordhoff, S., Winter, J. de, Payre, W., van Arem, B., & Happee, R. (2019). What impressions do users have after a ride in an automated shuttle? An interview study. Transportation Research Part F: Traffic Psychology and Behaviour, 63, 252–269. https://doi.org/10.1016/j.trf.2019.04.009

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. Advances in Health Sciences Education : Theory and Practice, 15(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Oña, J. de (2021). Understanding the mediator role of satisfaction in public transport: A cross-country analysis. Transport Policy, 100, 129–149. https://doi.org/10.1016/j.tranpol.2020.09.011

On-Road Automated Driving (ORAD) committee. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. 400 Commonwealth Drive, Warrendale, PA, United States. SAE International.

Pearl, J. (2012). The causal foundations of structural equation modeling. In Handbook of structural equation modeling (pp. 68–91). The Guilford Press.

Petrović, Đ., Mijailović, R., & Pešić, D. (2020). Traffic Accidents with Autonomous Vehicles: Type of Collisions, Manoeuvres and Errors of Conventional Vehicles’ Drivers. Transportation Research Procedia, 45, 161–168. https://doi.org/10.1016/j.trpro.2020.03.003

Pituch, K. A., & Stevens, J. (2016). Applied multivariate statistics for the social sciences: Analyses with SAS and IBM’s SPSS (6th edition). Routledge/Taylor & Francis Group.

Rafael, S., Correia, L. P., Lopes, D., Bandeira, J., Coelho, M. C., Andrade, M., Borrego, C., & Miranda, A. I. (2020). Autonomous vehicles opportunities for cities air quality. Science of the Total Environment, 712, 136546. https://doi.org/10.1016/j.scitotenv.2020.136546

Robert M. Schindler, & Barbara Bickart. (2005). Published Word of Mouth: Referable, Consumer-Generated Information on the Internet. https://www.researchgate.net/publication/-252952246_Published_Word_of_Mouth_Referable_Consumer_Generated_Information_on_the_Internet

Roche-Cerasi, I. (2019). Public acceptance of driverless shuttles in Norway. Transportation Research Part F: Traffic Psychology and Behaviour, 66, 162–183. https://doi.org/10.1016/j.trf.2019.09.002

Salonen, A. O. (2018). Passenger's subjective traffic safety, in-vehicle security and emergency management in the driverless shuttle bus in Finland. Transport Policy, 61, 106–110. https://doi.org/10.1016/j.tranpol.2017.10.011

Sina Nordhoff, Bart Van Arem, Natasha Merat, Ruth Madigan, & Riender Happee (2017). User Acceptance of Driverless Shuttles Running in an Open and Mixed Traffic Environment. In 12th ITS European Congress. https://www.researchgate.net/publication/317497564_User_Acceptance_of_Driverless_Shuttles_Running_in_an_Open_and_Mixed_Traffic_Environment

Site, P. D., Filippi, F., & Giustiniani, G. (2011). Users' preferences towards innovative and conventional public transport. Procedia - Social and Behavioral Sciences, 20, 906–915. https://doi.org/10.1016/j.sbspro.2011.08.099

Smolnicki, P. M., & Sołtys, J. (2016). Driverless Mobility: The Impact on Metropolitan Spatial Structures. Procedia Engineering, 161, 2184–2190. https://doi.org/10.1016/j.proeng.2016.08.813

SPACE. (2021, June 8). AutoMOST. https://space.uitp.org/initiatives/automost-pilot-malaga-spain

Stephanie Glen. (2016). Kaiser-Meyer-Olkin (KMO) Test for Sampling Adequacy. StatisticsHowTo.com. https://www.statisticshowto.com/kaiser-meyer-olkin/

Stocker, A., & Shaheen, S. (2017). Shared automated vehicles: Review of business models (International Transport Forum Discussion Paper 2017-09). Paris: Organisation for Eco-nomic Co-operation and Development (OECD), International Transport Forum. https://www.econstor.eu/handle/10419/194044

Tomás, R. F., Fernandes, P., Macedo, E., Bandeira, J. M., & Coelho, M. C. (2020). Assessing the emission impacts of autonomous vehicles on metropolitan freeways. Transportation Research Procedia, 47, 617–624. https://doi.org/10.1016/j.trpro.2020.03.139

TRATON GROUP. (2019). TRATON – Full-length autonomous buses. TRATON GROUP. https://traton.com/en/newsroom/current-topics/Nobina-and-Scania-.html

Venkatesh, Morris, & Davis (2003). User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27(3), 425. https://doi.org/10.2307/30036540

Volvo Buses Global. (2019). NTU Singapore and Volvo unveil world's first full size, autonomous electric bus [Press release]. https://www.volvobuses.com/en/news/2019/mar/volvo-and-singapore-university-ntu-unveil-world-first-full-size-autonomous-electric-bus.html

W.H. Whyte (1954). The web of word of mouth. Fortune, 50(5), 140–143. https://www.research-gate. net/publication/283439981_The_web_of_word_of_mouth

Wu, X., Cao, J., & Huting, J. (2018). Using three-factor theory to identify improvement priorities for express and local bus services: An application of regression with dummy variables in the Twin Cities. Transportation Research Part a Policy and Practice, 113, 184–196. https://doi.org/ 10.1016/j.tra.2018.04.003

YUTONG. (2020). Yutong 5G-Enabled Intelligent Mobility Solution. YUTONG Autonomous driving. https://en.yutong.com/technology/autonomous-driving/

Zhang, W., Guhathakurta, S., Fang, J., & Zhang, G. (2015). Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach. Sustainable Cities and Society, 19, 34–45. https://doi.org/10.1016/j.scs.2015.07.006