Transportation Research Part C 188 (2026) 105646

Contents lists available at ScienceDirect

Transportation Research Part C

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

Use of travel time in a shared automated vehicle for work and leisure: Results from a field experiment with a Wizard-of-Oz simulator-on-wheels vehicle

Maryna Öztürker [a,*], Sina Nordhoff [b], Sascha Hoogendoorn-Lanser [c], Bart van Arem [a], Gonçalo Homem de Almeida Correia [a]

[a] Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands
[b] EV Research Center, Institute of Transportation Studies, University of California, 1605 Tilia Street, Davis, CA 95616, USA
[c] Mobility Innovation Centre Delft, TU Delft Campus, Building 26, Van der Burghweg 1, 2628CS Delft, the Netherlands

A R T I C L E   I N F O

Keywords:
Shared automated vehicles
Non-driving related tasks
Stated choice experiment
Experimental configurations
Wizard-of-Oz simulator-on-wheels vehicle
Ride experience

A B S T R A C T

Shared automated vehicles (SAVs) have the potential to transform travel by enabling users to engage in non-driving-related tasks (NDRTs), enhancing productivity and travel satisfaction. To explore this potential, we conducted a field experiment using a Wizard-of-Oz simulator-on-wheels replicating SAV services in urban areas. The study examined how engagement in work and leisure NDRTs influenced attitudes, preferences, and associated values of travel time (VoTTs) for SAVs versus conventional transport modes (public transport (PT), cars, and bicycles). A total of 104 participants completed two test rides while engaging in work and leisure activities, with engagement levels captured via video recordings. Results showed that travel costs for SAVs were perceived as less negative than those of PT and cars, and that participants preferring work over leisure in SAVs developed a more positive perception of travel time in them post-test. In contrast, full concentration on NDRTs during test rides increased the disutility of travel time of the car alternative. Pre-test results indicated that SAVs had the highest VoTTs compared to cars and PT. However, after the rides, VoTTs for SAVs decreased when used for work-related activities, underscoring their advantage for productivity-focused travel. For cars, the ability to fully concentrate on NDRTs increased VoTTs, reflecting heightened expectations of comfort and productivity. These findings highlight SAVs' potential to enhance travel productivity, but also show how experience with NDRTs reshapes conventional modes perceptions. Finally, the experiment demonstrated the relevance of the Wizard-of-Oz approach for simulating realistic SAV experiences, with 74% of participants believing the setup was genuine.

1. Introduction

> *"In the future, driving a vehicle might become like horse riding. Just a hobby"*
> Anonymous participant 1

* Corresponding author at: Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P. O. Box 5048, 2600 GA Delft, the Netherlands.
E-mail address: m.ozturker@tudelft.nl (M. Öztürker).

https://doi.org/10.1016/j.trc.2026.105646
Received 17 February 2025; Received in revised form 15 January 2026; Accepted 16 March 2026
Available online 21 April 2026

"Staying curious helps to stop anxiety when daring to ride in a self-driving vehicle"
Anonymous participant 2

In contemporary society, commuting constitutes a significant part of individuals' daily routines, often seen as an unproductive yet unavoidable activity. For instance, in the Netherlands, average one-way door-to-door commutes last approximately 80 min by train, 45 min by urban public transport (bus, tram, or metro), and 25 min by car or bicycle (CBS, 2023c). To mitigate time loss and improve commute experience, passengers of trains, urban public transport (PT), and cars commonly engage in non-driving-related tasks (NDRTs), also known as travel-based multitasking. Common activities are working, reading, socialising with fellow passengers, or relaxing, which allow travellers to use their time productively or enjoyably (Keseru et al., 2020; Rizki et al., 2021; Singleton, 2020; Sun & Wong, 2022).

Innovations in transport technology, such as automated vehicles (AVs), present opportunities to transform commuting by enhancing both productivity and travel satisfaction. Driverless AVs operating at SAE Levels 4 and 5 eliminate the need for active driver engagement, allowing travellers to dedicate their commute entirely to NDRTs, such as sleeping, working, or socialising (SAE, 2021; Wadud & Huda, 2019). However, the deployment of AVs comes with challenges. Simply replacing conventional vehicles with privately owned AVs may exacerbate urban congestion and increase travel demand, undermining potential benefits (Barreto et al., 2022; Meyer et al., 2017; Milakis et al., 2017). A more sustainable approach involves deploying AVs as car- or ride-sharing services in urban areas and integrating them with PT systems (Acheampong et al., 2021; Bala et al., 2023; Carrese et al., 2023; Fan et al., 2023).

Shared automated vehicle (SAV) services, exemplified by companies like Waymo, Baidu, Didi, and Apollo, are actively testing their technology through pilot programs in cities across China and the USA. While these programs demonstrate advancements in vehicle automation, empirical research is needed to assess how effectively travellers can utilise SAV commute time for NDRTs and whether such activities enhance travel productivity and satisfaction, which can become a catalyst for SAV adoption. This is particularly relevant given that, in conventional transport modes like trains and urban PT, engaging in NDRTs has been shown to improve travel satisfaction by reducing the perceived disutility and value of travel time (VoTT) (Kouwenhoven & de Jong, 2018; Malokin et al., 2021; Molin et al., 2020; Varghese & Jana, 2018; Wardman et al., 2020). However, despite well-documented evidence of travel-based multitasking in traditional modes, research on NDRT engagement in SAVs remains limited. Understanding these dynamics is essential for fostering SAV adoption and ensuring that AVs as car- and ride-sharing services contribute to sustainable urban mobility rather than exacerbating congestion.

Prior research on travel time use for NDRTs in AVs has shown that potential users express preferences for activities such as sleeping, watching movies, reading, or messaging, with preferences varying based on the AV's level of automation and whether it is privately owned or shared (Kyriakidis et al., 2015; Bansal et al., 2016). Travel distance further shapes NDRT preferences, with shorter trips favouring smartphone use and longer trips encouraging resting or working, especially among AV-oriented users (Lee et al., 2021). While AVs are unlikely to replace fixed-location activities (i.e., activities previously done at home or the office), they expand the range of NDRTs performed during travel, particularly among individuals with higher income and education levels (Pudāne et al., 2018; 2019; 2021).

Vehicle interior design also plays a role in facilitating NDRTs. AVs with work-oriented interior redesigns showed lower VoTTs compared to those redesigned for leisure. However, contrary to expectations, AVs with redesigned interiors had higher VoTTs compared to chauffeur-driven vehicles with the same interior modifications (Correia et al., 2019).

Few empirical studies have directly assessed individuals' ability to engage in NDRTs inside AVs, as most research focuses on responses to takeover requests during NDRTs rather than the tasks themselves (de Winter et al., 2014; Naujoks et al., 2018; Shahini & Zahabi, 2022). For instance, Ko & Ji (2018) measured the flow experience of reading and watching videos in Level 3 AVs in a driving simulator experiment. Their findings suggest that moderate task difficulty can induce a state of flow, a mental state characterised by complete immersion and focus. Similarly, Klingegård et al. (2020) demonstrated through a Wizard-of-Oz experiment that participants could perform mentally demanding tasks in Level 4 AVs on highways as effectively as in an office environment.

While prior research provides valuable insights, it addresses either behavioural or experiential aspects of travel time use for NDRTs in AVs, leaving the interplay between NDRT engagement, mode preferences, and changes in the VoTT in the context of SAVs largely unexplored.

Three key research gaps emerge. First, survey-based studies, often using stated choice experiments, rely on hypothetical scenarios, potentially introducing hypothetical bias (Haghani et al., 2021). This bias may cause participants to overestimate their preferences and attitudes toward AVs, limiting insights into realistic engagement with NDRTs in SAVs. Second, empirical studies focus on the feasibility of engagement in NDRTs under real-world conditions, in which factors such as vehicle dynamics, automated driving-induced motion sickness, physical efforts and cognitive demands affect travellers' experiences (Bellem et al., 2018; Cornet et al., 2022; Elbanhawi et al., 2015; Singleton, 2019). While these studies benefit from the collection of physiological data – such as heart rate, electrodermal activity, electroencephalography, glance behaviour, and head posture estimation – to provide objective insights into the ability to engage in NDRTs inside AVs, they do not capture how such engagement affects travellers' attitudes, preferences, and VoTT for AVs. Third, prior research has largely focused on AVs in general, often in comparison with conventional private vehicles, rather than specifically investigating how SAVs, as a distinct transport mode, facilitate engagement in NDRTs. To the best of our knowledge, only one study has explicitly examined NDRTs in the context of SAVs, finding that car-sharing AVs were generally preferred over ride-sharing AVs and public transport. Preferences were strongly influenced by the ability to perform NDRTs, such as reading, using social media, and gaming, though certain activities, like writing, negatively impacted SAV choice (Hamadneh & Esztergár-Kiss, 2022). However, this study also relied on hypothetical scenarios, highlighting the ongoing gap between perceived and real-world behaviours.

These gaps underscore the need for robust empirical research to accurately assess how the use of travel time for NDRTs in SAVs influences travellers' attitudes, preferences, and VoTT under real-world conditions. Addressing these gaps is essential to inform strategies for integrating SAVs into urban PT systems and maximising their potential benefits on the individual and societal levels.

To address the gaps, this study employs a Wizard-of-Oz (WoZ) experiment, a controlled alternative setup for real-world testing that simulates SAV operations while minimising safety risks and navigating regulatory restrictions, as pilot programs with SAVs are not yet permitted in most countries, particularly across Europe. Participants will experience both work- and leisure-related NDRTs while travelling in urban areas, providing a realistic context for evaluating their attitudes and preferences toward SAVs. This division into work- and leisure-related activities is adopted to simplify participants' understanding, as it aligns with familiar, everyday scenarios and reflects common distinctions in travel-based multitasking (Cornet et al., 2022; Keseru & Macharis, 2018). Changes in participants' attitudes, preferences, and associated VoTTs before and after the WoZ experiment are assessed through stated choice (SC) surveys, capturing the influence of experiencing NDRTs in SAVs. SAVs are evaluated in comparison with other travel alternatives that include not only private conventional cars but also urban PT modes, such as buses and trams, as well as bicycles, which are widely used in the Netherlands. Finally, the study measures the extent of participants' engagement in NDRTs during SAV rides, offering further insights into how these activities affect travel behaviour and perceptions of SAVs.

Building on this experimental framework, the present study advances the understanding of behavioural and experiential aspects of travel time use for NDRTs in SAVs through the following contributions:

- Development and deployment of a WoZ experiment that allows participants to experience work- and leisure-related NDRTs while travelling in urban areas.
- Investigation of how using travel time for work- and leisure-related NDRTs in SAVs influences users' attitudes, preferences, and associated VoTTs compared to conventional transport modes, including conventional cars, urban PT (bus or tram), and bicycles, as evaluated through a stated preference experiment.
- Examination of consistency in users' attitudes, preferences, and associated VoTTs for SAVs, conventional cars, urban PT, and bicycles before and after the field experiment, capturing the effect of experiencing NDRTs.
- Measurement of user engagement levels in work- and leisure-related NDRTs during SAV rides to assess the depth and quality of involvement.

The remainder of this paper is organised as follows. In Section 2, based on the review of experimental configurations used in AV studies, categorised as in-lab, on-road, and mixed experiments, we adopt a WoZ simulator-on-wheels vehicle setup designed to explore productive travel time use in SAVs. In Section 3, we describe the four-step methodology, which combines subjective data collection methods – such as pre- and post-test SC surveys and semi-structured interviews – with physiological data collection methods aimed at capturing NDRTs' engagement during test rides in an experimental SAV. The modelling approach and discussion of results are given in Sections 4 and 5, respectively. In Section 6, the main conclusions of the present study are outlined.

2. Experimental configurations for studying travel time use in automated vehicles

Real-life AV experiences are essential for addressing the limitations of studies that rely on mental images and perceptions of automated driving technology. These experiences enhance research validity by providing accurate assessments of user attitudes and preferences. However, conducting experiments in complex traffic conditions presents challenges due to safety risks and legal restrictions (Etzioni et al., 2021; Farooq et al., 2018; Greifenstein, 2024; Lukovics et al., 2023; Zou et al., 2021). Despite advancements in commercial AV operations in some regions, pilot programs involving AVs remain restricted or entirely prohibited in most countries, particularly across Europe.

To overcome these challenges, researchers have developed alternative experimental configurations categorised as in-lab, on-road, and mixed setups. These configurations differ in terms of perceived realism, safety, costs, and data collection feasibility – key factors for ensuring the validity and practicality of experiments. This section reviews these setups, highlighting their application in studying NDRTs. Based on this review, we adopt a WoZ simulator-on-wheels vehicle to explore the impact of travel time use on attitudes and preferences for SAVs.

2.1. Experimental configurations: in-lab, on-road, and mixed

Experimental configurations for studying AV-related phenomena can be broadly divided into in-lab, on-road, and mixed setups, each offering distinct advantages and limitations.

In-lab experimental configurations rely on traditional driving simulators and virtual reality (VR) simulators to recreate driving environments within controlled settings. Traditional simulators typically include a driver's seat, steering wheel, and pedals, with fidelity ranging from simple setups using basic screens to high-fidelity systems mounted on motion platforms that replicate vehicle dynamics (de Winter et al., 2014; Ko & Ji, 2018; Minhas et al., 2020).

VR simulators enhance immersion through head-mounted displays (HMDs) or Cave Automatic Virtual Environments (CAVEs), where a virtual environment is projected onto surrounding walls (Ejichukwu et al., 2024; Kettle & Lee, 2022; Riegler et al., 2021). These setups are cost-effective, safe, and enable streamlined data collection across controlled scenarios. While CAVE systems represent a more advanced form of immersive simulation, they, and VR setups more broadly, often lack the full range of multisensory and physical stimuli present in real-world travel, such as vehicle motion, road surface vibration, dynamic lighting, and noise variability. As a result, such setups may fall short in replicating the unpredictability and associated experience of real-world traffic conditions, thereby limiting ecological validity.

On-road experimental configurations take place on test tracks or public roads. Test tracks can range from simple closed circuits to advanced proving ground facilities simulating urban environments with diverse road types and intersections (Hartwich et al., 2019; Lucovics et al., 2023; Shi & Bengler, 2022; Chen et al., 2020; Yang et al., 2021). They offer a safer and more predictable alternative to real traffic. These controlled conditions also facilitate reliable and consistent data collection, making test tracks particularly effective for studying specific driving scenarios.

Public road experiments, often conducted on highways, provide the highest realism by exposing participants to genuine traffic conditions. However, they involve greater safety risks, logistical challenges, and difficulties in collecting consistent data due to the dynamic and unpredictable nature of real-world environments (Dillmann et al., 2023; Feys et al., 2021; Naujoks et al., 2016; Noble et al., 2021; Solís-Marcos et al., 2018).

Vehicles in on-road experiments include AVs with lower automation levels (Levels 1–3) (Hartwich et al., 2019; Liu et al., 2019; 2021) and WoZ vehicles simulating higher levels of automation (Levels 3–5), where concealed drivers mimic automation while maintaining safety (Ekman et al., 2019; Naujoks et al., 2019; Osz et al., 2018).

Mixed setups combine elements of in-lab and on-road experiments, offering hybrid environments. Examples include the WoZ simulator-on-wheels vehicle, which integrates real-world driving with simulator-like displays (Baltodano et al., 2015; Detjen et al., 2020), and the WoZ VR simulator-on-wheels, where participants experience a virtual overlay of real-world traffic through HMDs while riding in a vehicle (Zou et al., 2021). These configurations balance realism, safety, cost, and data collection feasibility, making them promising for AV research.

2.2. Exploring NDRTs across experimental configurations

In-lab and on-road experiments are widely used to study engagement in NDRTs during automated driving, whereas mixed experimental setups remain limited overall and particularly for studying NDRT engagement.

In-lab simulator studies primarily examine drivers' reaction times, workload, and situational awareness during transitions from NDRTs to manual control (Naujoks et al., 2018; Riegler et al., 2021; Shahini & Zahabi, 2022; Gerber et al., 2020; Li et al., 2020). However, research on NDRT performance in AVs using VR simulators is limited due to the technology's low fidelity in replicating the complexity of NDRTs. Additionally, HMDs restrict natural movement, making them unsuitable for tasks requiring a full range of motion or intricate physical actions (Riegler et al., 2021).

On-road experiments explore how NDRTs affect drivers' ability to monitor lower-level automation (Levels 1–2) and take over control in Level 3 AVs. These studies also examine how automation level and driver experience influence NDRT performance (Naujoks et al., 2016; Naujoks et al., 2019; Noble et al., 2021; Shi & Bengler, 2022; Solís-Marcos et al., 2018). While offering high realism, these setups face challenges in ensuring participant safety and consistent data collection.

Mixed setups, such as the WoZ simulator-on-wheels and the WoZ VR simulator-on-wheels, allow participants to experience dynamic urban environments. However, NDRTs have only been studied in the WoZ simulator-on-wheels, where participants' self-selection of tasks during rides was monitored (Detjen et al., 2020). The WoZ VR simulator-on-wheels, while promising, inherits limitations of HMD-based VR simulators, including restricted movement and reduced feasibility for tasks requiring full motion (Zou et al., 2021).

Across all setups, NDRTs typically involve everyday activities (e.g., reading) or standardised tasks (e.g., n-Back tasks), with tasks typically performed on handheld devices, mounted displays, or head-up displays. However, most studies focus on functional performance metrics, neglecting deeper aspects of the feasibility of NDRT engagement – a gap highlighted earlier in Section 1. Exceptions include Ko and Ji (2018), who measured flow experience in simulators, and Klingegård et al. (2020), who evaluated NDRT engagement using a WoZ vehicle in a Level 4 highway scenario.

2.3. Evaluating and selecting the experimental configuration

To identify the most suitable experimental setup for studying travel time use for NDRTs in SAVs operating in urban environments, this evaluation considers four key factors: perceived realism, safety risks, costs, and data collection feasibility. These factors are derived from reviewed studies and additional sources reflecting key elements that influence the validity, practicality, and replicability

Table 1
Comparison of experimental configurations across key factors.

Experimental configuration	Key factors Perceived realism[a]	Safety risks[b]	Costs[c]	Data collection feasibility[d]
In-lab	Low	High	Low	High
On-road (test track)	Moderate	Moderate	Moderate	Moderate
On-road (public road)	High	Low	High	Low
Mixed	High	Moderate	Moderate-High	Moderate

[a] (de Winter et al., 2012; Detjen et al., 2020; Eriksson et al., 2017; Hartwich et al., 2019; Riegler et al., 2021; Weidner et al., 2017; Zou et al., 2021).
[b] (de Winter et al., 2012; Detjen et al., 2020; Eriksson et al., 2017; Riegler et al., 2021; Zou et al., 2021).
[c] (Baltodano et al., 2015; Eriksson et al., 2017; Riegler et al., 2021; Zou et al., 2021).
[d] (de Winter et al., 2012; Eriksson et al., 2017; Riegler et al., 2021).

of experimental research (Baltodano et al., 2015; de Winter et al., 2012; Detjen et al., 2020; Eriksson et al., 2017; Hartwich et al., 2019; Riegler et al., 2021; Weidner et al., 2017; Zou et al., 2021). Table 1 summarises the differences between the experimental configurations based on these factors.

A mixed experimental setup emerges as the optimal choice for this study, offering a balance between realism, safety, cost, and feasibility. Among mixed configurations, the WoZ simulator-on-wheels vehicle provides distinct advantages by combining the high realism of on-road experiments with the safety and control of in-lab setups. This configuration enables participants to engage in realistic NDRT scenarios while experiencing urban traffic conditions.

The WoZ simulator-on-wheels addresses the ecological validity limitations of in-lab studies and mitigates the safety risks of public road experiments. Its ability to simulate higher levels of automation (Levels 3–4) further enhances its relevance for future SAV research. While mixed setups involve higher costs and logistical complexities, their benefits outweigh these challenges, making the WoZ simulator-on-wheels a robust platform for exploring how productive travel time use influences user attitudes and preferences toward SAVs.

3. Methodology

This field study employs a convergent mixed-method approach that combines subjective and physiological data collection methods to explore the impact of using travel time for work- and leisure-related NDRTs in SAVs on users' attitudes, preferences, and associated VoTTs compared to conventional transport modes, including cars, urban PT (bus or tram), and bicycles.

The core element of the study involves two test rides in a WoZ simulator-on-wheels vehicle, where participants engage in both work and leisure-related NDRTs while travelling through urban areas. Subjective data is gathered through pre- and post-test online surveys, complemented by semi-structured interviews. Physiological data is gathered during the test rides via facial video recordings to quantify the extent of participants' engagement in NDRTs.

The methodology is structured around a four-step process, outlined in Fig. 1. The subsequent sections provide a detailed overview of this approach.

3.1. Procedure

Experimental vehicle. The test rides were conducted using a Wizard-of-Oz (WoZ) simulator-on-wheels vehicle to simulate a Level 5 SAV (MICD, 2022), a setup similar to Detjen et al. (2020). The vehicle, a Nissan e-NV200 Evalia van provided by LeasePlan, was modified to create the illusion of full automation while being manually driven by a "safety driver". To maintain this illusion, a wall

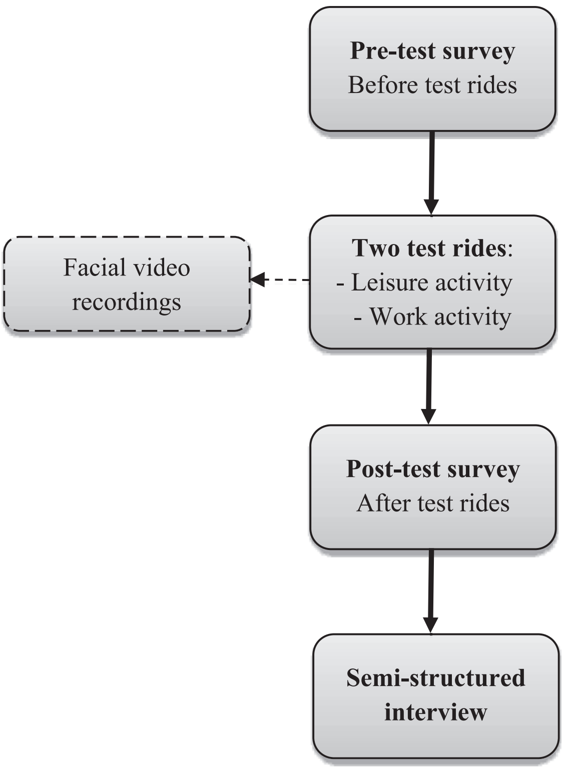

Fig. 1. Four-step methodological approach.

partition was installed behind the first row of seats (Figs. 2, 3), fully enclosing the passenger compartment. Inside, three OLED screens simulated a driver's seat view by displaying real-time footage of the driving environment, captured via cameras mounted on the front and side windshields. Participants had no access to driving controls, such as the steering wheel or pedals, reinforcing the perception of a fully autonomous ride.

The passenger compartment was designed to provide privacy and comfort, allowing one participant at a time to engage in work- and leisure-related NDRTs in a distraction-free setting. The seating arrangement ensured that the participant sat in the rear seat, directly behind the partition, with an unobstructed view of the simulated road environment.

To ensure the participant's safety and comfort, an emergency stop button was installed on the left side of the seat, allowing participants to terminate the experiment immediately if they experienced any discomfort or negative effects.

A camera mounted above the central front display recorded participant behaviour during the test rides, providing data on NDRT engagement levels. Facial video recordings, captured by this camera, were used to assess the extent of participants' engagement in NDRTs. This video data was chosen as a non-invasive method to allow participants to freely use handheld devices and materials during the test rides without interference. Inspired by camera-based driver monitoring systems (Pech et al., 2019; Seaman et al., 2022; van Gent et al., 2017), head orientation – particularly a downward position while focusing on handheld devices or reading materials – served as a proxy for attention allocation and engagement level.

Pre-test. Given that only one participant at a time could be in the experimental WoZ vehicle, participants were invited to book a timeslot via a link provided in the recruitment advertisement. The advertisement included a brief study description, outlining its stages and time commitment, which consisted of approximately 20 min for the pre-test survey and 1.5 h on the day of the experiment.

One week before their scheduled appointment, participants received an email containing a link to the pre-test survey, which they were required to complete at home. The survey assessed their initial attitudes and preferences toward SAVs compared to conventional transport modes.

Before accessing the pre-test survey, participants were required to sign a consent form, which, along with the standard description of the research procedure and data handling practices, also included a confidentiality clause. This clause prohibited participants from discussing the experimental setup until the study was completed, preventing them from revealing the true nature of the experimental vehicle to others.

Along with the survey, participants received a detailed description of the study procedure, outlining the steps they would follow on

a) Exterior view


b) Interior set-up

c) Passenger during a test ride

Fig. 2. Wizard-of-Oz simulator-on-wheels vehicle.

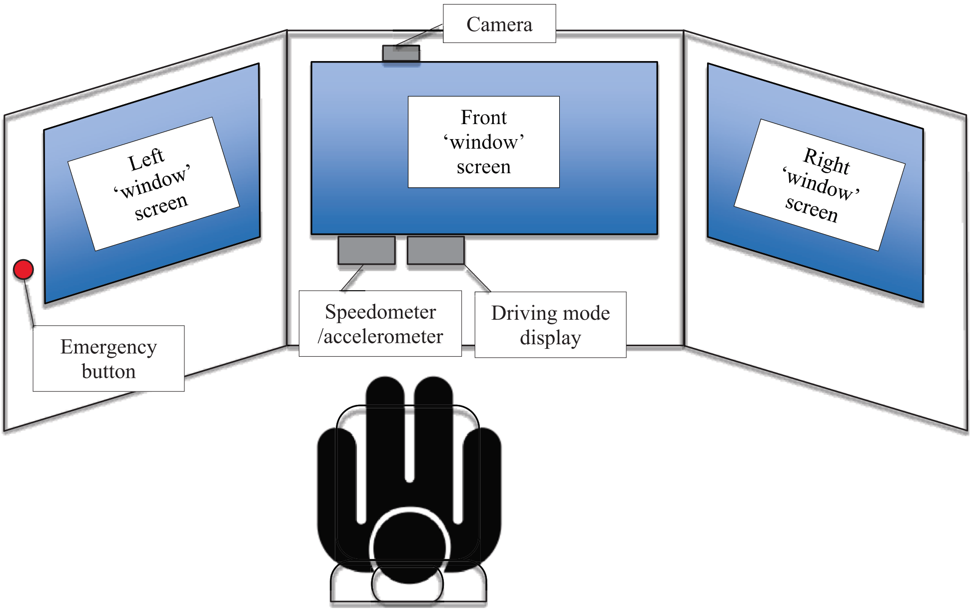

Fig. 3. Schematic setup of the passenger's compartment.

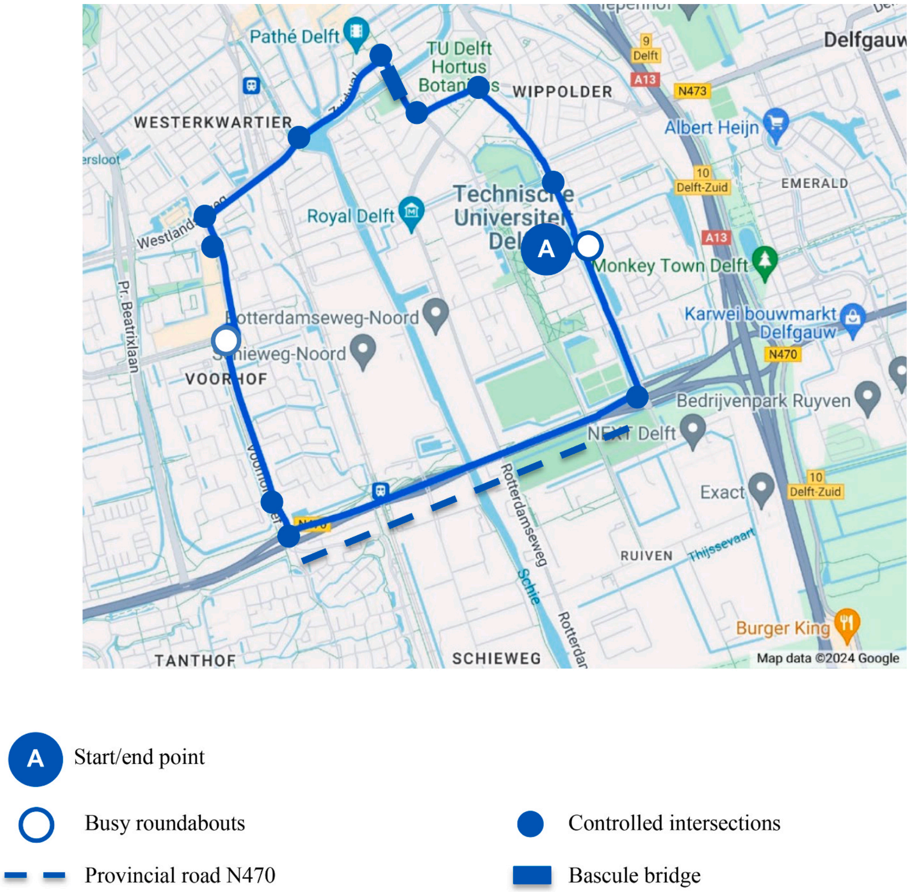

Fig. 4. Test rides route in Delft (approximately 20 min).

the day of the experiment. This included a 10-minute introduction, followed by two test rides of approximately 20 min each. After each ride, participants would complete a post-test survey and take part in a short interview during a 20-minute break to discuss their experience. Additionally, they were provided with a list of materials required for work- and leisure-related NDRTs to ensure they were adequately prepared.

The work- and leisure-related NDRTs were carefully selected to be actively engaging, cognitively demanding, and intrinsically motivating, based on criteria from Cornet et al. (2022) and Keseru & Macharis (2018). To ensure ecological validity and accommodate diverse participant interests, the selection of specific tasks was left to participants. They were provided with an example list of possible tasks prior to the test day (Appendix A) and encouraged to choose their own tasks similar to those examples, specifically, activities they would normally perform during travel or idle time.

Given the 20-minute test ride duration, suitable work-related tasks included planning an agenda, responding to emails, or reading. Participants were asked to bring their laptops and prepare work-related tasks in advance; for unemployed or retired participants, laptops were provided on-site along with example tasks such as writing a letter or email, creating a grocery list, or planning a weekly agenda. Leisure activities included reading paper-based materials or using a smartphone, with participants either bringing their own or selecting from provided options. This approach aimed to maintain high engagement, allow personal relevance, and simulate realistic travel time use during SAV rides.

Experimental day. During the introduction session, participants were first introduced to the experimental scenario, in which they were asked to imagine calling an SAV via an app, which would then arrive at their doorstep to take them to their destination.

To ensure their comfort and safety, participants were informed that they could terminate the experiment at any time by pressing an emergency button (Fig. 3) if they felt uncomfortable. They were also advised to use the initial minutes of the ride to acclimate before beginning their assigned tasks. The "safety driver" was introduced as being present to oversee the vehicle's operation and ensure a safe driving experience.

Following the introduction, each participant took two test rides in the WoZ simulator-on-wheels vehicle, engaging in both work- and leisure-related NDRTs. To control for potential order effects, participants were randomly assigned to start with either work-related or leisure-related tasks, ensuring a balanced experimental design.

The vehicle followed a 7.2 km route through Delft, the Netherlands, covering an urban driving environment with a mix of roundabouts, regulated intersections, busy city streets (max 50 km/h), a bascule bridge, and a provincial road (max 80 km/h) (Fig. 4). Test rides were conducted on weekdays during daylight hours to ensure optimal lighting for NDRT engagement and to maintain consistent traffic conditions by avoiding peak hours. A team of seven "safety drivers" was instructed to drive calmly and cautiously, reflecting the safe and precautionary driving style of an AV.

After each test ride, participants took a 20-minute break, during which they completed a post-test survey and participated in semi-structured interviews to discuss their experience.

At the end of the 1.5-hour session, researchers disclosed the Wizard-of-Oz setup, clarifying that the vehicle was manually driven. Participants were also reminded of the confidentiality agreement, ensuring they would not share details about the experiment until the study period was completed.

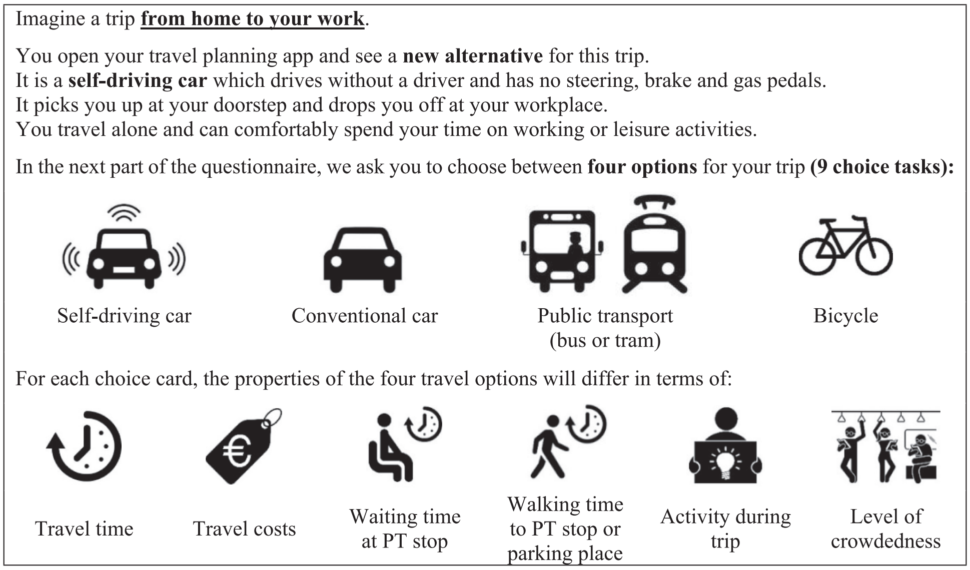

Fig. 5. Instructions shown to participants before starting the SC experiment.

3.2. Instrument

3.2.1. Pre- and post-test surveys

The pre-test survey conducted before the test rides encompassed (a) questions regarding the participants' current travel behaviour, (b) an SC experiment on mode choice, (c) indicator statements to gauge attitudes towards SAVs, and (d) questions about the participants' socio-economic background.

In the post-test survey after each test ride, participants were asked to repeat the SC experiment and reassess the indicator statements. The surveys were available in Dutch and English.

The opening section of the pre-test survey contained questions about the respondents' current travel behaviour, such as current travel mode, trip frequency and duration. The final section included questions regarding the respondents' socio-economic background, covering aspects such as gender, age, educational level, occupation, annual gross household income, possession of a driving license and PT pass, ownership of various types of vehicles, history of traffic accidents, presence of mobility restrictions or motion sickness, use of car- or ride-sharing services, and experience with AVs.

The SC experiment was designed to explore the preferences for SAVs, considering the possibility of performing NDRTs when using this car-sharing service. Before presenting the choice sets, participants were provided with detailed instructions about the SC experiment (Fig. 5).

The choice sets included four labelled alternatives: an SAV (referred to as a self-driving car for clarity), a conventional car, PT (bus or tram), and a bicycle. Participants were asked to choose their preferred mode of transportation under scenarios where in-vehicle travel time, travel costs, waiting time at the PT stop or doorstep, and walking time to the PT stop or parking place were provided. For the SAV alternative, the attribute "Activity during trip" had two levels: working and leisure. Additionally, the level of crowdedness was included as an attribute of the PT (bus or tram) alternative.

The attribute levels were primarily determined based on data from trip-planning applications for mid-sized cities in the Netherlands (Table 2). The travel costs for the conventional car alternative were calculated for middle-class vehicles (i.e., vehicles typically priced and categorised between economy and luxury, offering moderate features and performance) and included expenses for fuel, insurance, maintenance, and taxes, excluding parking costs (Nibud, 2023). The travel costs for SAVs offering car-sharing services were assumed to be similar to those for the conventional car alternative.

Based on the participants' indicated occupation, the purpose of the trip was specified in the choice sets in the direction from home to the workplace for employed or self-employed individuals, a place of study for students, or any frequently visited location for unemployed or retired individuals.

Each respondent was presented with a total of nine hypothetical choice tasks, repeated in pre-test and post-test surveys (see example of a choice task in Fig. 6). The choice tasks were generated using Ngene software through an orthogonal design, resulting in 36 choice sets divided into four blocks (ChoiceMetrics, 2018).

Moreover, in the pre-test and post-test surveys, participants rated their agreement with 18 attitudinal statements on a 7-point Likert scale, representing six psychological constructs: enjoyment of AVs and the ride experience, perceived safety and trust in their technological capabilities, intention to use SAVs, service quality, and perception of work and leisure activities. The last two constructs specifically assess the importance of using travel time for work or leisure, as well as the comfort and ability to concentrate on these activities during the ride (Table 3).

3.2.2. Semi-structured interviews

To gain a deeper understanding of participants' perceptions of SAVs, their experiences of riding in the Wizard-of-Oz simulator-on-wheels vehicle and engaging in work- and leisure-related NDRTs, we conducted semi-structured interviews, which took place with participants during the breaks after the first and the second test rides, after participants filled in the post-test surveys.

The semi-structured interviews began with questions about participants' motivations for joining the experiment, as well as their knowledge of and interest in automated driving technology. Throughout the conversation, the researcher had a subtle objective to determine whether participants genuinely believed in the experimental setup. In other words, whether they thought they were truly riding in a self-driving vehicle. Consequently, the conversation was steered toward participants' perceptions of safety, their trust level in the automated driving capabilities of SAVs, and how these factors influenced their ability to engage in the NDRTs during test rides.

Table 2
Attribute levels.

Attributes and attribute levels	Alternative 1 Self-driving car	Alternative 2 Conventional car	Alternative 3 Public transport (bus or tram)	Alternative 4 Bicycle
Travel time (min)	15 / 25 / 35	15 / 25 / 35	15 / 25 / 35	20 / 30 / 40
Travel costs (€)	3.0 / 4.0 / 5.0	3.0 / 4.0 / 5.0	2.4 / 3.4 / 4.4	–
Waiting time (min)	2 / 5 / 8	–	2 / 5 / 8	–
Walking time (min)	–	2 / 4 / 6	4 / 7 / 10	–
Activity	(0) Leisure (1) Work	–	–	–
Crowdedness	–	–	(0) Not crowded (1) Light crowdedness (2) Crowded	–

Card 1. Which mode of transport do you prefer to travel from your home to work in this situation?

	Self-driving car	Conventional car	Public transport (bus or tram)	Bicycle
Travel time	**35 min**	**35 min**	**25 min**	**20 min**
Travel costs	**4 euro**	**3 euro**	**2.4 euro**	-
Waiting time	**5 min**	-	**2 min**	-
Walking time	-	**4 min**	**10 min**	-
Activity	**Working**	-	-	-
Crowdedness	-	-	**Light crowdedness**	-

Fig. 6. Example of a choice task (for employed or self-employed participants).

Participants were also asked to compare the driving style of the automated system with that of a human driver. Following the second test ride, the focus of the interviews went on contrasting their experiences of engaging in work- versus leisure-related tasks.

3.3. Participants

A total of 104 participants took part in the five-week experiment conducted between November and December 2022. This main sample was recruited from the Delft panel with support from the Delft municipality. As an incentive, participants had the opportunity to win one of ten €50 gift cards.

Because the main sample consisted of volunteers from an existing panel, there was a risk of self-selection bias: these participants might have been more curious or enthusiastic about experiencing SAVs than the general population. To address this concern, we recruited a control group of 35 participants through social media, personal networks, and paper-based advertisements, following recommendations from Feys et al. (2021) and Liu et al. (2019). Unlike the main sample, participation in the control group was voluntary and non-incentivised. These participants completed a pre-test survey that excluded the experimental scenario and served as a baseline to check for the differences in choices.

Although the recruitment strategies differed (panel vs open call), both groups included individuals from Delft and surrounding areas with diverse backgrounds in age, gender, education, and travel patterns. Appendix B summarises the socio-economic characteristics of both groups (Table B.1), mode choices by trip purpose (Table B.2), and additional participant details (Table B.3).

4. Modelling approach

4.1. Joint model specification

In this study, we employed a mixed logit model with panel effects, jointly estimated on pre- and post-test datasets, following methodologies from Jensen et al. (2013) and González et al. (2016). This approach allowed us to compare participants' preferences before and after experiencing work and leisure activities during test rides in the experimental vehicle while accounting for correlation in repeated choices from the same individuals. To account for potential heteroscedasticity between the pre- and post-test datasets, a scale parameter was included in the joint model formulation. This is a common approach in stated preference studies to address differences in unobserved utility variance, which can arise due to variations in geographic regions, datasets, or, in our study, the two

Table 3
Attitudinal indicators and underlying psychological constructs.

	Indicators*	Source	Likert scale
Enjoyment of AVs			
S1	I like self-driving cars	Adapted from Öztürker et al. (2022)	1 = dislike extremely, 7 = like extremely
S2	I think that a ride in a self-driving car is enjoyable (A ride in a self-driving car was enjoyable)	Adapted from Nordhoff et al. (2018a)	1 = strongly disagree, 7 = strongly agree
S3	I think that a ride in a self-driving car is stressful (A ride in a self-driving car was stressful)	Adapted from Yap et al. (2016)	
Perceived safety and trust			
S4	I trust that a system can drive a self-driving car with no assistance from me	Adapted from Yap et al. (2016)	1 = strongly disagree, 7 = strongly agree
S5	I dislike that I don't have control of how the car drives	Adapted from Kyriakidis et al. (2015)	
S6	I can entrust the safety of a close family member to a self-driving car		
S7	I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)		
Work activity			
S8	It is important for me to use my travel time productively when I'm riding in a self-driving car (I would use my travel time productively when I ride in a self-driving car)	Created for this study	1 = strongly disagree, 7 = strongly agree
S9	I think I will be (I was) able to concentrate on working in a self-driving car		
S10	I think it will be (It was) comfortable to work in a self-driving car		
S11	I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)	Adapted from Kyriakidis et al. (2015)	
Leisure activity			
S12	I think I will be (I was) able to concentrate on my leisure activities in a self-driving car	Created for this study	1 = strongly disagree, 7 = strongly agree
S13	I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car		
S8 and S11		See above	
Intention to use shared automated vehicles			
S14	I like that an electric self-driving car does not produce pollutant emissions	Adapted from Yap et al. (2016)	1 = strongly disagree, 7 = strongly agree
S15	In the future, I will use self-driving cars for my daily trips	Adapted from Nordhoff et al. (2018b)	
S16	I think that a ride in a self-driving car saves time (would save my time)	Adapted from Kyriakidis et al. (2015)	
Service quality			
S17	I am afraid that there will be no car available when I request one	Adapted from Yap et al. (2016)	1 = strongly disagree, 7 = strongly agree
S18	I am worried that the car is not clean after its previous use		

* The statements were modified for the post-test survey, see in parentheses.

survey waves (Train, 2009; Jensen et al., 2013; González et al., 2016). More flexible specifications, such as random scale mixed logit or latent class models (Hess & Train, 2017), can also capture variation in scale across individuals or subgroups, but they require larger samples and add estimation complexity. Given that our joint model incorporated a panel structure and was estimated on the available sample (n = 104), a fixed scale adjustment was selected as the most suitable choice for our objective of assessing before-after changes in preferences following experimental SAV exposure. The model was estimated using the PandasBiogeme software package (Bierlaire, 2023).

In the joint mixed logit model with panel effects, $U_{in}^{pre-test}$ and $U_{in}^{post-test}$ represent utilities that individual n assigns to alternative i before and after the test rides, respectively (Eq. (1)).

$$U_{in}^{pre-test} = \sum_k \beta_{ik}^{pre-test} \bullet x_{ikn} + \sum_l \beta_{il}^{pre-test} \bullet x_{ln}^{pre-test} + \sum_s \beta_{is}^{pre-test} \bullet x_{sn} + \gamma_{in}^{pre-test} + \varepsilon_{in}^{pre-test} \tag{1}$$

$$U_{in}^{post-test} = \mu\left(\sum_k \beta_{ik}^{post-test} \bullet x_{ikn} + \sum_l \beta_{il}^{post-test} \bullet x_{ln}^{post-test} + \sum_s \beta_{is}^{post-test} \bullet x_{sn} + \gamma_{in}^{post-test} + \varepsilon_{in}^{post-test}\right)$$

where:

x_{ikn} and β_{ik} is the vector of instrumental variables and their estimated parameters;
x_{ln} and β_{il} is the vector of latent variables and their estimated parameters;
x_{sn} and β_{is} is the vector of socio-economic variables and their estimated parameters;
γ_{in} is a normally distributed error component capturing panel effects (mean 0, standard deviation σ);
ε_{in} is an independent and identically distributed (i.i.d.) extreme value type 1 error term;
μ is a scale parameter to normalise error variances across pre- and post-test data.

The first component of the utility function includes instrumental variables x_{ikn} that represent observable attributes of the travel alternatives in the SC experiment, such as in-vehicle travel time, travel costs, waiting time at the PT stop or doorstep, walking time to the PT stop or parking place, activity during the trip, and level of crowdedness (see Table 2). Level of crowdedness for the public PT (bus or tram) alternative was dummy-coded with three levels: "Not crowded" (reference category), "Light crowdedness", and "Crowded". For the SAV alternative, the variable "Activity during trip" enters the model as an interaction term with in-vehicle travel time, travel costs, and waiting time at the doorstep, allowing the effect of on-board activity to vary depending on these trip attributes.

The second component comprises latent variables x_{ln} that capture unobservable psychological constructs derived from attitudinal statements on perceptions, concerns, and beliefs about automated transport. Because latent variables cannot be directly observed, confirmatory factor analysis (CFA) was applied to a set of 18 attitudinal indicators (see Table 3). Unlike exploratory factor analysis (EFA), which uncovers latent structures without prior assumptions, CFA is suited to testing hypothesised structures (Brown, 2015). In our case, the same attitudinal indicators were collected in two waves (pre- and post-test), and we required a consistent factor structure to meaningfully compare latent constructs across them. EFA tends to yield different factor solutions across datasets, particularly when applied separately to each wave, which complicates longitudinal comparisons. We therefore used CFA to confirm the hypothesised structure and assessed measurement invariance via multi-group CFA to confirm stability across waves (Brown et al., 2017).

After confirming the latent variable structure through CFA, we specified structural equations to integrate the constructs x_{ln} into the discrete choice model, estimated separately for the pre-test and post-test stages S, as the attitudes of a participant n could change after engaging in work and leisure activities during test rides

$$x_{ln}^S = k^S + \sum_s \varphi_s^S \bullet x_{sn} + \omega_n^S \tag{2}$$

where k is the intersept, x_{sn} and φ_s is the vector of socio-economic variables and their estimated parameters and ω_n is the error term with zero mean and standard deviation σ_ω.

The relationship between each latent variable x_{ln} and $r = 1 \cdots R$ indicators I_{rn} associated with it was modelled through measurement equations

$$I_{rn}^S = k_r^S + \sum_l \alpha_r^S \bullet x_{ln}^S + \lambda_{rn}^S \tag{3}$$

where k_r is the intersect, x_{ln} is the latent variable, α_r is the factor loading linking the latent variable to indicator r and λ_{rn} is the error term with zero mean and standard deviation σ_λ.

Given the complexity of estimating the model with several latent variables, we integrated the latent variables into the discrete choice model sequentially, as this approach helps to manage the computational challenges involved. This sequential integration approach is recognised in the literature as a practical, though less statistically efficient, alternative to full-information estimation. In complex hybrid models, sequential estimation allows latent constructs to be incorporated stepwise without overburdening the estimation routine, while still yielding meaningful and statistically significant results (Bahamonde-Birke & Ortúzar, 2014; Johansson et al., 2006).

The third component accounts for individual socio-economic characteristics x_{sn} such as age, gender, and education, as well as mobility-related attributes like car ownership and current main transport mode (see Tables B.1 – 3). By including these variables, the model captures systematic heterogeneity in mode choice that cannot be explained by instrumental attributes or latent attitudes. To incorporate categorical socio-economic variables into the model, we applied dummy coding. The variable categories and their corresponding reference categories used in the coding scheme are detailed in Tables B.1 – 3.

In the utility functions (Eq. (1)), the parameters β_{ik}, β_{il}, β_{is} of instrumental, latent, and socio-economic variables, respectively, were initially estimated as specific to the pre- and post-test experimental stages. This was done to determine whether work and leisure activities during the test rides influenced participants' preferences. When no significant differences were found between pre- and post-test preferences, generic parameters were introduced.

The utility functions specified in Eq. (1) include two types of error terms that reflect different sources of unobserved variation. First, the term γ_{in} is a normally distributed error component (mean zero, standard deviation σ) that accounts for panel effects by capturing unobserved individual-specific variation across repeated choices. The term ε_{in} is an independent and identically distributed extreme value type 1 (EV1) error, capturing random noise at the observation level. To account for potential differences in error variance between the pre- and post-test datasets, a scale parameter μ is introduced. This normalises the variance across waves, allowing a joint model to be estimated (Jensen et al., 2013; González et al., 2016).

Finally, two interaction terms were included in the post-test utility functions to account for participants' test-ride experiences. The first captures the engagement levels in work- and leisure-related NDRTs derived from video analysis and dummy-coded into three levels: “Full concentration”, “Partial concentration” (either on work or leisure activities), and “No concentration” (reference category). The second interaction accounts for whether participants believed in the experimental setup, based on post-ride interviews, and was coded as “Believed”, “Hesitant”, and “Did not believe” (reference). These variables interact with instrumental, latent, and socio-economic variables in the post-test model to assess whether engagement and belief moderated mode preferences. As a result, Eq. (1) was extended to incorporate these interaction effects in the post-test utility specification, where C stands for “Concentration” and B stands for “Belief”:

$$U_{in}^{pre-test} = \sum_k \beta_{ik}^{pre-test} \bullet x_{ikn} + \sum_l \beta_{il}^{pre-test} \bullet x_{ln}^{pre-test} + \sum_s \beta_{is}^{pre-test} \bullet x_{sn} + \gamma_{in}^{pre-test} + \varepsilon_{in}^{pre-test} \tag{4}$$

$$\begin{aligned} U_{in}^{post-test} = \mu \Bigg(& \sum_k \beta_{ik}^{post-test} \bullet x_{ikn} + \sum_l \beta_{il}^{post-test} \bullet x_{ln}^{post-test} + \sum_s \beta_{is}^{post-test} \bullet x_{sn} \\ & + \sum_{conc \bullet k} \delta_{ik}^{post-test} \bullet x_{ikn} \bullet C_{in} + \sum_{conc \bullet l} \delta_{il}^{post-test} \bullet x_{ln} \bullet C_{in} + \sum_{conc \bullet s} \delta_{is}^{post-test} \bullet x_{sn} \bullet C_{in} \\ & + \sum_{belief \bullet k} \delta_{ik}^{post-test} \bullet x_{ikn} \bullet B_{in} + \sum_{belief \bullet l} \delta_{il}^{post-test} \bullet x_{ln} \bullet B_{in} + \sum_{belief \bullet s} \delta_{is}^{post-test} \bullet x_{sn} \bullet B_{in} + \gamma_{in}^{post-test} + \varepsilon_{in}^{post-test} \Bigg) \end{aligned}$$

Finally, a similar modelling approach (based on the utility specification in Equation (1)) was used to compare the pre-test preferences of the main sample with those of the control group.

4.2. Modelling steps

The modelling process followed a stepwise approach (Fig. 7).

Step 1. Half of the participants began with work-related activities, and the other half with leisure activities. To test whether responses differed systematically depending on activity order, we estimated a joint multinomial logit (MNL) model with instrumental variables.

Full taste homogeneity was tested under the assumption that preferences were identical across all instrumental parameters ($\beta_{post-first\ ride} = \beta_{post-second\ ride}$), with potential differences existing only in scale ($\mu_{post-first\ ride} \neq \mu_{post-second\ ride}$). The correctness of the joint model specification was verified using the likelihood ratio test (de Dios Ortúzar & Willumsen, 2011):

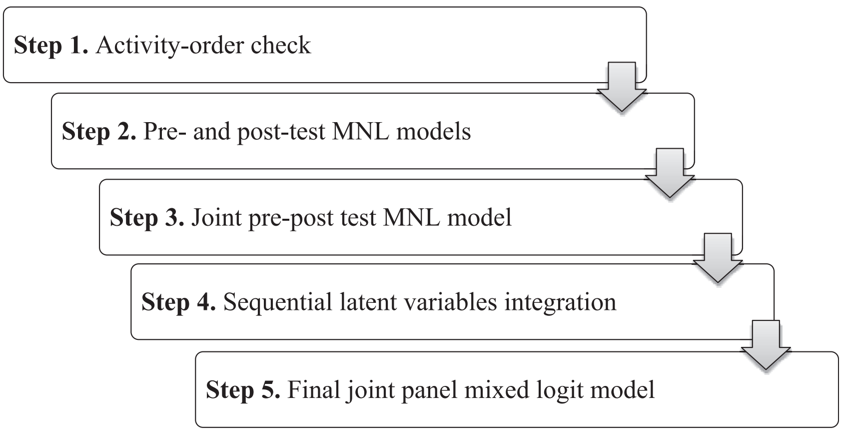

Fig. 7. Modelling steps.

$$LRS = -2\{L(\beta_j) - L(\beta)\} \quad (5)$$

where $L(\beta_j)$ is the final log-likelihood of the joint model, and $L(\beta)$ is the sum of the final log-likelihoods of the independently estimated post-first and post-second ride models. The independent models produced a combined log-likelihood of $L(\beta) = -1911.3333$ with a total of 32 parameters (16 each), while the joint model with 16 generic parameters and one scale parameter yielded $L(\beta_j) = -1913.849$. The likelihood ratio statistic (5.0314) was far below the χ^2 critical value of 24.996 for 15 degrees of freedom at a 95% significance level. We therefore conclude that responses exhibit full taste homogeneity between the post-first and post-second test rides. As a result, subsequent analysis focused on responses collected after the second test ride, when participants had experienced both activities.

Step 2. We then estimated separate MNL models for the pre-test and post-test datasets using instrumental and socio-economic variables ($MNL_{pre\text{-}test}$ – 21 parameters, $LL_{pre\text{-}test} = -957.4702$; $MNL_{post\text{-}test}$ – 26 parameters, $LL_{pre\text{-}test} = -856.3805$). At this stage,

Table 4
Scores on attitudinal indicators in the main sample before and after test rides and control group.

	Indicators	Scores (mean /standard deviation) Control group	Pre-test	Post-test	Wilcoxon signed-rank test statistics between: Control group – Pre-test	Pre-test – Post-test
Enjoyment of AVs						
S1	I like self-driving cars	4.89 (1.409)	5.16 (1.158)	5.80 (0.907)	Z = −0.052	Z = −4.875**
S2	I think that a ride in a self-driving car is enjoyable (A ride in a self-driving car was enjoyable)	5.29 (1.202)	5.42 (1.103)	6.03 (1.000)	Z = −0.650	Z = −4.381**
S3	I think that a ride in a self-driving car is stressful (A ride in a self-driving car was stressful)	4.17 (1.248)	4.74 (1.386)	5.85 (1.399)	Z = −1.527	Z = −5.730**
Perceived safety and trust						
S4	I trust that a system can drive a self-driving car with no assistance from me	4.57 (1.596)	5.07 (1.457))	5.76 (1.170)	Z = −0.724	Z = −5.079**
S5	I dislike that I don't have control of how the car drives (reversed)	3.83 (1.671) (1.671)	4.01 (1.721)	4.98 (1.625)	Z = −0.399	Z = −5.157**
S6	I can entrust the safety of a close family member to a self-driving car	4.34 (1.392) (1.392)	4.63 (1.515)	5.41 (1.319)	Z = −0.217	Z = −5.303**
S7	I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)	4.40 (1.499)	5.0 (1.231)	6.02 (0.750)	Z = −1.143	Z = −6.702**
Work activity						
S8	It is important for me to use my travel time productively when I'm riding in a self-driving car (I would use my travel time productively when I ride in a self-driving car)	4.94 (1.514)	4.44 (1.717)	5.28 (1.523)	Z = −1.928	Z = −4.098**
S9	I think I will be (I was) able to concentrate on working in a self-driving car	4.09 (1.853)	4.55 (1.461)	5.49 (1.435)	Z = −0.663	Z = −4.864**
S10	I think it will be (It was) comfortable to work in a self-driving car	4.40 (1.701)	4.77 (1.331)	5.05 (1.523)	Z = −0.877	Z = −1.591
S11	I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)	5.26 (1.094)	5.13 (1.089)	5.57 (1.197)	Z = −0.915	Z = −2.867**
Leisure activity						
S12	I think I will be (I was) able to concentrate on my leisure activities in a self-driving car	5.11 (1.491)	5.05 (1.361)	5.56 (1.44)	Z = −0.816	Z = −3.297**
S13	I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car	5.23 (1.285)	5.01 (1.296)	5.59 (1.312)	Z = −1.628	Z = −3.549**
S8 and S11		See above				
Intention to use shared automated vehicles						
S14	I like that an electric self-driving car does not produce pollutant emissions	5.69 (0.963)	6.28 (1.101)	6.28 (1.065)	Z = −1.446	Z = −0.071
S15	In the future, I will use self-driving cars for my daily trips	4.89 (1.676)	4.48 (1.707)	5.08 (1.419)	Z = −2.076*	Z = −3.903**
S16	I think that a ride in a self-driving car saves time (would save my time)	4.83 (1.524)	4.62 (1.382)	5.33 (1.310)	Z = −1.219	Z = −4.956**
Service quality						
S17	I am afraid that there will be no car available when I request one (reversed)	3.71 (1.152)	3.63 (1.436)	3.96 (1.481)	Z = −0.248	Z = −2.325*
S18	I am worried that the car is not clean after its previous use (reversed)	4.26 (1.358)	4.09 (1.596)	4.41 (1.629)	Z = −0.167	Z = −2.846**

The difference between the two scores is significant based on Wilcoxon signed-rank test statistics (IBM, 2017);
** at a 99% level; * at a 95% level.

Table 5
Results of confirmatory factor analysis: main sample (pre-test).

Factor	Indicators (attitudinal statements)	Factor loadings	Reliability Internal reliability (Cronbach's alpha)	Composite reliability	Convergent validity (average variance extracted)
		$RMSEA_{pre\text{-}test} = 0.081$; $CFI_{pre\text{-}test} = 0.965$; $TLI_{pre\text{-}test} = 0.942$			
Factor 1. Perceived safety, trust and enjoyment of AVs	S1. I like self-driving cars	0.65	0.86	0.87	0.63
	S4. I trust that a system can drive a self-driving car with no assistance from me	0.87			
	S6. I can entrust the safety of a close family member to a self-driving car	0.80			
	S7. I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)	0.83			
Factor 2. Work activity	S9. I think I will be (I was) able to concentrate on working in a self-driving car	0.70	0.76	0.79	0.66
	S10. I think it will be (It was) comfortable to work in a self-driving car	0.91			
Factor 3. Leisure activity	S11. I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)	0.80	0.81	0.76	0.52
	S12. I think I will be (I was) able to concentrate on my leisure activities in a self-driving car	0.62			
	S13. I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car	0.73			
Factor 4. Intention to use SAVs	S15. In the future, I will use self-driving cars for my daily trips	0.82	0.66	0.68	0.52
	S16. I think that a ride in a self-driving car saves time (would save my time)	0.61			

non-significant parameters were excluded at a 10% significance level, given the intermediate nature of this step in the modelling process.

Step 3. The two MNL models were subsequently combined into a joint pre-/post-test MNL under the assumption of partial taste homogeneity. Full homogeneity was considered unlikely, so candidate parameters were identified using covariance/correlation analysis of parameter pairs with *t*-values below 1.96 (95% significance) (de Dios Ortúzar & Willumsen, 2011). PandasBiogeme (Bierlaire, 2023) reports this analysis, and 8 generic parameters and a scale parameter were introduced into the joint model (MNL_{joint} – 38 parameters, $LL_{joint} = -1818.304$). As in Step 1, the model specification was tested using the likelihood ratio test (Eq. (5)). At the 95% significance level, the likelihood ratio statistic (8.9066) was lower than the χ^2 critical value of 16.919 (9 df), confirming correct specification of the joint model.

Step 4. After validating the latent constructs through CFA, we conducted multi-group CFA to test measurement invariance, confirming that the same attitudinal constructs could be meaningfully compared across both waves (see Section 5.1.1). This provided the foundation for the structural part of the MIMIC model, in which socio-demographic variables were specified as causes of the latent constructs. To manage computational complexity, we first pre-estimated the latent variables model (see Section 5.1.2) and then introduced the latent variables sequentially into the joint model, ensuring that each construct could be incorporated without overburdening the estimation routine.

The order of inclusion follows theoretical priority and ensures comparability across time points. Constructs most central to our research focus, namely perceived work activity and perceived leisure activity, are integrated first, followed by the remaining constructs in decreasing order of measurement reliability as indicated by the CFA. For each construct, we first added the pre-test latent variable, immediately followed by its post-test counterpart. This ordering allowed a balanced assessment of attitudinal changes across the two survey waves. While sequential estimation is less efficient than simultaneous full-information approaches, it is widely applied in hybrid choice modelling as a pragmatic solution when estimation complexity prevents joint estimation (Bahamonde-Birke & Ortúzar, 2014; Johansson et al., 2006).

Step 5. In the final step, we estimated a joint mixed logit model that incorporated an error component, γ_{in}, to account for panel effects. This component is normally distributed with mean zero and standard deviation σ, and captures unobserved individual-specific variation across repeated tasks. The results of the final joint mixed logit model with panel effects are presented in Table 10 and discussed in Section 5.4, 5.5.

We repeated Steps 2 – 5 using only the pre-test responses, comparing the main sample with the control group. The corresponding results of the joint mixed logit model with panel effects are reported in Table 13 and further examined in Section 5.6.

5. Results and discussion

5.1. Results of the latent variables model

5.1.1. Validating the latent constructs through CFA

Latent variables in the second term of the utility functions (Eq. (1)) represent unobservable psychological constructs that form a latent variables model. The hypothesised structure of the latent variables model consists of 6 latent variables, namely 'Enjoyment of AVs', 'Perceived safety and trust', 'Work activity', 'Leisure activity', 'Intention to use shared automated vehicles' and 'Service quality' (see Table 3).

Before validating the structure of the latent model, we first examined the average scores that participants assigned to the indicators (Table 4). Both the main sample and the control group showed generally positive attitudes toward SAVs and the potential for productive use of travel time while travelling in them. Notably, the scores in the main sample increased significantly after the test rides (see Pre-test – Post-test column in Table 4), whereas no significant differences were observed between the pre-test scores and those given by the control group on nearly all indicators (see Control group – Pre-test column in Table 4).

Afterwards, CFA was performed separately on the pre-test, post-test, and control group datasets to validate the latent variables model. In other words, it was used to assess the relationships between attitudinal indicator statements and their corresponding latent constructs.

In the main sample, the first run of CFA with the hypothesised model structure revealed low and non-significant standardised loadings for all indicators of the "Service quality" latent construct. Due to this poor measurement performance and limited contribution to the structural model, the construct was excluded from the final estimation (Appendix C, Tables C.1 and C.4). Additionally, the factors "Enjoyment of AVs" and "Perceived safety and trust" were highly correlated when checked for discriminant validity using the Fornell-Larcker criterion, leading to their combination into a single factor (Appendix C, Tables C.2, C.3 and C.5, C.6). The final model consisted of four latent constructs: "Perceived safety, trust, and enjoyment of AVs", "Work activity", "Leisure activity", and "Intention to use SAVs". Each latent construct was represented by at least two indicators, and factor loadings were estimated using maximum likelihood estimation with the lavaan R package (Rosseel, 2012). The CFA results for the pre-test and post-test datasets are detailed in Tables 5 and 6.

The goodness-of-fit indices, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA), for both the pre-test and post-test models indicated an acceptable fit (Tables 5, 6). RMSEA reflects how well the model fits the population covariance matrix, with values below 0.08 indicating acceptable fit and values below 0.05 considered good. CFI and TLI assess model fit relative to a baseline model, with values above 0.90 generally indicating acceptable fit and values above 0.95 considered good (Brown, 2015). These indices were selected because they are widely accepted in CFA as indicators of model adequacy.

Indicators with insignificant or weak factor loadings (<0.6) were removed from both models. While a loading of 0.7 is typically recommended, loadings of 0.6 or above are considered acceptable when supported by theory and model fit (Hair et al., 2021). All

Table 6
Results of confirmatory factor analysis: main sample (post-test).

Factor	Indicators (attitudinal statements)	Factor loadings	Reliability Internal reliability (Cronbach's alpha)	Composite reliability	Convergent validity (average variance extracted)
		$RMSEA_{post\text{-}test} = 0.075$; $CFI_{post\text{-}test} = 0.972$; $TLI_{post\text{-}test} = 0.954$			
Factor 1. Perceived safety, trust and enjoyment of AVs	S1. I like self-driving cars	0.73	0.84	0.83	0.54
	S4. I trust that a system can drive a self-driving car with no assistance from me	0.77			
	S6. I can entrust the safety of a close family member to a self-driving car	0.67			
	S7. I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)	0.78			
Factor 2. Work activity	S9. I think I will be (I was) able to concentrate on working in a self-driving car	0.85	0.89	0.89	0.80
	S10. I think it will be (It was) comfortable to work in a self-driving car	0.93			
Factor 3. Leisure activity	S11. I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)	0.76	0.82	0.76	0.52
	S12. I think I will be (I was) able to concentrate on my leisure activities in a self-driving car	0.66			
	S13. I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car	0.74			
Factor 4. Intention to use SAVs	S15. In the future, I will use self-driving cars for my daily trips	0.74	0.77	0.77	0.63
	S16. I think that a ride in a self-driving car saves time (would save my time)	0.84			

remaining loadings were significant ($p < 0.001$) and strong indicators of their respective latent constructs.

To further assess the reliability and validity of the latent variables, Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) were calculated (Tables 5, 6). Thresholds of 0.70 for Cronbach's alpha and CR, and 0.50 for AVE, are commonly used to indicate acceptable internal consistency and convergent validity (Hair et al., 2021; Cheung et al., 2024). Both pre-test and post-test models met these criteria, with Cronbach's alpha values exceeding 0.70, except for the "Intention to Use SAVs" factor in the pre-test model, which was kept for consistency with the post-test model. CR values also surpassed the 0.70 threshold, and AVE values confirmed convergent validity.

A multi-group confirmatory factor analysis was conducted in lavaan R package (Rosseel, 2012) to ensure model consistency across pre-test and post-test data (Brown et al., 2017). The results confirmed good model fit for both datasets, supporting measurement invariance across groups. This confirmation established that the latent constructs were stable across waves, thereby justifying their use in the structural part of the MIMIC model (Section 5.1.2). Multi-group CFA indicated good fit (CFI = 0.966; RMSEA = 0.081), and invariance tests (metric: ΔCFI = –0.003, ΔRMSEA = –0.001, scalar: ΔCFI = –0.009, ΔRMSEA = +0.013) showed $\Delta \text{CFI} \leq 0.01$ and $\Delta \text{RMSEA} \leq 0.015$, confirming the equivalence of constructs across waves.

Although the pre-test and control groups did not differ significantly in their average responses to attitudinal indicators, multi-group CFA revealed that the underlying factor structure was not comparable across the two groups (Table 7). Specifically, the configural model for the control group showed poor fit (CFI = 0.796, RMSEA = 0.178, SRMR = 0.100), indicating that participants organised the items differently into latent constructs (see Appendix C, Tables C.7-9 for model development stages). As outlined in Section 4.1, meaningful comparison of latent constructs across groups requires a consistent factor structure; without this, estimates of latent means or relationships are not valid. To avoid drawing misleading conclusions from non-equivalent latent variables, we therefore excluded the latent constructs from the joint model.

5.1.2. Latent variables model

Following the specification in Equation (2), each latent construct was regressed on socio-economic characteristics as part of the MIMIC model, namely gender, age, education level, occupation, and income. In this framework, the structural equations (Eq. (2)) capture the influence of socio-demographic causes on the latent constructs, while the measurement equations (Eq. (3)) link each construct to its observed indicators. The results of both the multiple indicators and multiple causes parts are presented in Table 8 (pre-test) and 9 (post-test). In what follows, we discuss how socio-economic characteristics shaped each of the four latent constructs, comparing pre-test results (Table 8) with post-test results (Table 9).

For the first factor, representing perceptions of safety, trust, and enjoyment of AVs, pre-test results suggest that women, the highly educated, and higher-income respondents held more positive perceptions, whereas older and employed participants expressed negative views. Post-test results reveal substantial shifts: the effect of gender and income turns negative, and the contribution of education strengthens considerably. The age penalty remains but diminishes, while the effect of employment is no longer significant. These results indicate that direct exposure reshaped perceptions in ways that attenuated or reversed several of the pre-test socio-demographic gradients.

The second factor, relating to the perceived suitability of working in SAVs, was initially evaluated more positively by women and higher-income respondents, while older participants and those in employment expressed more negative views. After the test rides, the gender advantage remained but was smaller, the negative age effect became stronger, the employment penalty decreased sharply, and the income effect disappeared altogether. These patterns suggest that experience with the vehicle reduced income-related differences but reinforced age-based scepticism about working during rides.

The third factor, reflecting leisure-related activities in SAVs, displayed a contrasting pattern. Before the rides, women, older

Table 7
Results of confirmatory factor analysis: control group.

Factor	Indicators (attitudinal statements)	Factor loadings	Reliability Internal reliability (Cronbach's alpha)	Composite reliability	Convergent validity (average variance extracted)
		$RMSEA_{control} = 0.122$; $CFI_{control} = 0.953$; $TLI_{control} = 0.924$			
Factor 1. Perceived safety, trust and enjoyment of AVs	S1. I like self-driving cars	0.7	0.77	0.79	0.57
	S2. I think that a ride in a self-driving car is enjoyable (A ride in a self-driving car was enjoyable)	0.95			
	S4. I trust that a system can drive a self-driving car with no assistance from me	0.56			
Factor 3. Leisure activity	S8. It is important for me to use my travel time productively when I'm riding in a self-driving car (I would use my travel time productively when I ride in a self-driving car)	0.67	0.84	0.86	0.60
	S11. I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)	0.72			
	S12. I think I will be (I was) able to concentrate on my leisure activities in a self-driving car	0.71			
	S13. I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car	0.97			

Table 8
Results of the latent variables model (pre-test).

Indicators / Predictors	Factor 1 Perceived safety, trust and enjoymentof AVs	Factor 2 Work activity	Factor 3 Leisure activity	Factor 4 Intention to use SAVs
Multiple indicators				
S1. I like self-driving cars	1			
S4. I trust that a system can drive a self-driving car with no assistance from me	−1.58***			
S6. I can entrust the safety of a close family member to a self-driving car	−1.45***			
S7. I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)	−0.776***			
S9. I think I will be (I was) able to concentrate on working in a self-driving car		1		
S10. I think it will be (It was) comfortable to work in a self-driving car		−0.897***		
S11. I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)			0.371***	
S12. I think I will be (I was) able to concentrate on my leisure activities in a self-driving car			1	
S13. I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car			0.943***	
S15. In the future, I will use self-driving cars for my daily trips				1
S16. I think that a ride in a self-driving car saves time (would save my time)				−0.69***
Multiple causes				
Gender: Female (Ref. – Male)	0.441***	0.642***	−0.553***	0.29***
Age: Old (above 50) (Ref. – Young)	−0.356***	−0.365***	−0.287***	−0.065
Education level: High (Ref. – Low)	0.11*	0.039	−0.149*	0.131*
Occupation: Employed (Ref. – Other)	−0.376***	−0.843***	0.841***	−0.62***
Income (gross annual per household): Above 50 k (Ref. – Below 50 k)	0.344***	0.183**	−0.288***	0.653***

*** significant at a 99% confidence interval; ** at 95%; * at 90%.

respondents, those with higher education, and higher-income participants reported less favourable attitudes, while the employed expressed stronger support. After exposure, the age effect flipped to a significant positive, female and income effects became non-significant, and the employment effect remained positive though weaker. Only the modest negative effect of education persisted. These findings suggest that riding experience was particularly effective in improving older participants' views of SAVs as spaces for leisure activities.

For the fourth factor, intention to use SAVs, pre-test results revealed stronger intentions among women, the highly educated, and higher-income respondents, while employment had a significant negative effect. After the rides, these relationships were reconfigured: higher income was linked to reduced intention, the effect of employment turned strongly positive, and the effects of gender and age disappeared. Education remained a small but consistent positive driver. Post-test intentions, therefore, became more closely aligned with being employed, while the pre-test prominence of income and gender weakened.

Taken together, the results show that direct exposure to SAVs reshaped how socio-economic characteristics influence perceptions and intentions. Before the rides, gender and income were strong predictors across several constructs, but their effects largely weakened or reversed afterwards. Education emerged as a consistently positive predictor of perceptions of safety, trust, and enjoyment as well as of intention to use SAVs. Employment, which was initially associated with more negative perceptions and intentions, became a positive predictor after the rides. Older participants remained sceptical about working in SAVs but became more favourable toward their leisure use. Overall, these patterns suggest that riding experience reduces pre-existing socio-economic divides, with education and employment emerging as the most important post-test determinants across the four latent constructs.

5.2. Quantifying the extent of engagement in NDRTs during test rides

To analyse participants' engagement in work- and leisure-related NDRTs during the test rides, facial video recordings from a camera installed above the central “front windshield” screen (Fig. 3) were processed using a computer vision algorithm estimating visual attention through head orientation (Nielsen, 2021). The algorithm calculated the percentage of time participants looked in five directions: forward, left, right, up, and down (Fig. 8), corresponding to the front screen, side “window” screens, the upper area of the front screen, and down at the task, respectively. Downward head orientation was used as a proxy for engagement, as gaze tracking was not feasible from the overhead camera position. When participants looked down at a task, their eyes were typically occluded from view, making head pose a practical and non-intrusive indicator of engagement for our experimental setting.

The use of head pose as a non-intrusive indicator of engagement is well established in naturalistic settings where direct eye tracking is either impractical or may alter participant behaviour. Hernandez et al. (2013) demonstrated that sustained head orientation toward

Table 9
Results of the latent variables model (post-test).

Indicators / Predictors	Factor 1 Perceived safety, trust and enjoymentof AVs	Factor 2 Work activity	Factor 3 Leisure activity	Factor 4 Intention to use SAVs
Multiple indicators				
S1. I like self-driving cars	1			
S4. I trust that a system can drive a self-driving car with no assistance from me	−0.395***			
S6. I can entrust the safety of a close family member to a self-driving car	−0.343***			
S7. I think that a ride in a self-driving car is safe (A ride in a self-driving car was safe)	−0.505***			
S9. I think I will be (I was) able to concentrate on working in a self-driving car		1		
S10. I think it will be (It was) comfortable to work in a self-driving car		−0.933***		
S11. I think that a ride in a self-driving car is comfortable (A ride in a self-driving car was comfortable)			0.468**	
S12. I think I will be (I was) able to concentrate on my leisure activities in a self-driving car			1	
S13. I think it will be (It was) comfortable to spend time for leisure activities in a self-driving car			0.799***	
S15. In the future, I will use self-driving cars for my daily trips				1
S16. I think that a ride in a self-driving car saves time (would save my time)				0.845***
Multiple causes				
Gender: Female (Ref. – Male)	−0.343***	0.213***	0.14	0.035
Age: Old (above 50) (Ref. – Young)	−0.271**	−0.466***	0.174**	−0.114
Education level: High (Ref. – Low)	0.698***	−0.032	−0.15**	0.131*
Occupation: Employed (Ref. – Other)	−0.049	−0.191***	0.401***	0.446***
Income (gross annual per household): Above 50 k (Ref. – Below 50 k)	−0.366***	0.0098	0.0468	−0.281***

*** significant at a 99% confidence interval; ** at 95%; * at 90%.

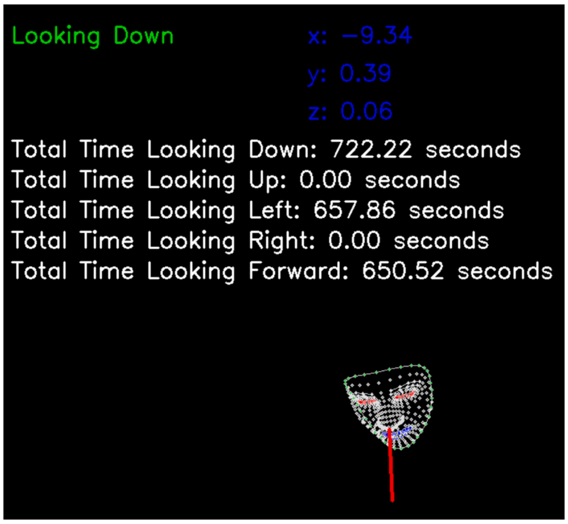

Fig. 8. Example of processed video output.

a screen can reliably distinguish between high and low engagement when monitoring television viewers, even without eye-movement data. Levordashka et al. (2023) similarly used head pose to capture audience attention during live events, showing its sensitivity to engagement patterns while allowing participants to behave naturally. In an educational context, Yu et al. (2021) applied head-pose estimation to monitor student engagement during lectures, further supporting its validity as an unobtrusive engagement metric in seated-task environments. These findings support our choice of head pose over more intrusive methods, which could have disrupted participants' self-selected NDRTs.

Participants who looked down for more than 70% of the ride were considered fully concentrated on the task. This threshold accommodates the preparatory phase at the start of the ride, when participants were instructed to first settle in the vehicle, observe their surroundings, and familiarise themselves with the interior before beginning their chosen task, as well as natural off-task glances to check the driving situation and posture adjustments. Our cut-off also aligns with previous engagement research: notably, Eriksson et al. (2018) reported that maintaining approximately 80% head orientation toward the relevant screen reflected an acceptable engagement level in an e-learning context, placing our 70% threshold within an empirically supported range while allowing for realistic off-task time in a naturalistic, in-ride setting.

Comparable thresholds appear in driver monitoring literature. In conventional driving, the NHTSA (2012) defines visual distraction when the Percent Road Center (PRC) falls below 60% in a given time window, and Kircher et al. (2009) use a similar 58% cut-off. Large-scale naturalistic studies report that attentive drivers keep their eyes on the road roughly 75–80% of the time (European Commission, 2015; 2022). In Level 3–4, on-road and simulator studies show that when drivers are fully engaged in a visual-manual NDRTs, they allocate 80–90% of gaze to the task (Britten, 2021; Klingegård et al., 2020).

Although our study was conducted in a vehicle simulating a Level 5 SAV, where occupants are not required to monitor the driving task, these established ranges support 70% as a conservative and empirically grounded cut-off. It ensures that "full concentration" reflects a clear majority of ride time dedicated to the task while preserving the naturalistic, non-intrusive character of the experiment.

Based on this criterion, 53.8% were fully engaged in both work and leisure activities, 10.6% focused only on work, 7.7% only on leisure, and 27.9% were not engaged in either. These results were dummy-coded into a parameter, "Concentration on NDRTs", with three levels: "Full concentration", "Partial concentration" (either on work or leisure activities), and "No concentration" (reference category).

This parameter was included as an interaction term in the discrete choice model (Eq. (4)) to explore whether preferences differed among participants based on their engagement levels. This provided insights into how varying engagement in NDRTs influenced mode choices and the associated VoTTs.

5.3. Belief in the experimental setup: results from semi-structured interviews

The semi-structured interviews provided valuable insights into participants' perceptions of SAVs, their trust in automated driving technology, and their experiences during the test rides. These interviews offered a deeper understanding of how participants engaged with the experimental setup and how they perceived the driving capabilities of the Wizard-of-Oz simulator-on-wheels vehicle.

The interview responses were manually annotated using predefined categories, including reasons for joining the study, knowledge and interest in automated driving technology, trust in SAVs and safety perception, comparison of driving styles (automated vs. human), and work- versus leisure-related NDRTs. A deductive content analysis approach (Elo & Kyngäs, 2008) was followed to classify participant responses within these categories and identify key patterns in their perceptions. Participants' initial belief in the authenticity of the experimental setup was also assessed, grouping them into three categories: those who believed the setup was genuine, those who did not, and those for whom it was unclear. After researchers disclosed the nature of the experimental setup, participants were asked to confirm whether they had initially believed the vehicle was self-driving.

The majority of participants cited curiosity as their primary motivation for joining the field study, followed by an interest in automated driving technology. Despite their curiosity, many were initially hesitant to fully trust the vehicle to drive autonomously. This hesitation was particularly evident among participants who appeared visibly nervous before the test rides. The presence of a "safety driver" reassured participants and contributed to their overall sense of security during the rides. Although most participants reported feeling safe, many were uncertain whether their confidence stemmed from the safety driver's presence or trust in the automated system itself.

Participants generally found the second test ride to be more relaxed, attributing this to their familiarity with the route and experimental settings from the first ride. However, no clear pattern emerged regarding whether work or leisure activities were easier or more difficult to perform, regardless of the order. This observation is further supported by the test model results (Subsection 4.2).

When asked whether they noticed a difference between the driving style of the automated system and a human driver, most participants reported little to no difference, primarily comparing it to their own driving behaviour. During the discussion, 58.7% of participants expressed amazement at the vehicle's automated driving capabilities, forming the initial group who believed the setup was genuine. In contrast, 1.9% immediately recognised that the vehicle was not a true AV, while 39.4% were uncertain about their belief in the setup until the experiment's nature was disclosed.

After researchers revealed that the vehicle was manually driven, 74% of participants confirmed they had genuinely believed they were riding in an AV, 18.3% remained uncertain until the disclosure, and 7.7% never believed the vehicle was fully automated. These findings indicate that the experimental setup created a realistic simulation of a self-driving vehicle for the majority of participants, reinforcing its relevance for studying user preferences and engagement in work- and leisure-related NDRTs during SAV rides. To examine whether these differences in belief affected participants' stated preferences, a "Belief in experimental setup" variable was derived from the post-ride interviews and included in the discrete choice model as an interaction term (Eq. (4)), dummy-coded as "Believed," "Hesitant," and "Did not believe" (reference category).

To further examine whether belief in the experimental setup varied systematically across participant groups, we conducted a series of Chi-square tests of independence. These analyses explored associations between belief categories and socio-demographic characteristics, as well as the order in which activities were performed during the rides. No significant associations were found with age ($\chi^2 = 2.475$, $p = 0.116$, Phi = 0.154), education level ($\chi^2 = 0.41$, $p = 0.522$, Phi = 0.063), occupation ($\chi^2 = 0.45$, $p = 0.502$, Phi = 0.066), or income ($\chi^2 = 1.522$, $p = 0.217$, Phi = 0.121). Similarly, belief did not differ depending on whether participants began with a work or leisure activity ($\chi^2 = 0.184$, $p = 0.668$, Phi = 0.042).

Notably, two quotes included at the beginning of this paper were drawn from these interviews, highlighting the participants' reflections on the experiment:

> *"In the future, driving a vehicle might become like horse riding. Just a hobby."*
> Anonymous participant 1
> *"Staying curious helps to stop anxiety when daring to ride in a self-driving vehicle."*
> Anonymous participant 2

5.4. *Results of the final joint discrete choice model*

The results of the joint panel mixed logit model are detailed in Table 10. Parameters are organised into three columns: the first two display stage-specific parameters for the pre- and post-test phases, which showed significant differences at the 95% level, reflecting the impact of engaging in work- and leisure-related NDRTs during the test rides. The third column includes generic parameters with no significant differences between the stages.

Parameter names include suffixes such as_SAV,_CAR,_PT, and_BIKE, which indicate the corresponding travel alternatives: shared automated vehicle, conventional car, public transport (bus or tram), and bicycle, respectively.

The following suffixes indicate reference to a specific instrumental, latent, or socio-economic parameter, with their meanings and category labels explained in the "Corresponding variable" column. For a complete overview of instrumental variables, see Table 2. Latent variables and their measurement indicators are provided in Tables 5 and 6, while socio-economic variables and their categorical coding are detailed in Tables B.1-B.3 in Appendix B.

The final model, comprising 51 parameters, was estimated using 1,000 Halton draws from a normal distribution, yielding stable and convergent results (Table 10). The adjusted Rho-squared value of 0.287, within the 0.2–0.4 range, indicates a strong fit to the data (Louviere et al., 2000). The model is intended to be exploratory and explanatory rather than predictive, aiming to uncover behavioural mechanisms rather than maximise out-of-sample forecasting performance. As shown in Table 11, the staged estimation process demonstrates a systematic improvement in model fit as additional behavioural structure is introduced: the inclusion of socio-economic variables increases the adjusted Rho-squared from 0.214 to 0.285, while the subsequent integration of latent variables yields a further improvement to 0.287, accompanied by a substantial gain in log-likelihood. While the Akaike Information Criterion (AIC) continues to decrease, indicating improved explanatory fit, the Bayesian Information Criterion (BIC) increases slightly in the final specification, reflecting its stronger penalty for model complexity. This trade-off is consistent with the exploratory nature of the model and suggests that the latent variables add explanatory insight at the cost of increased complexity, rather than indiscriminate overfitting. Additionally, the post-test scale parameter differs significantly from 1, indicating a 23% higher variance in post-test error terms relative to the pre-test data, thereby justifying the joint estimation approach with scale adjustment.

5.5. *Discussion*

When discussing the results in the following subsections, we focus on three key aspects that align with the main objectives of this study. First, we explain how participants' experience of engagement in work- and leisure-related NDRTs during the test rides affected their preferences for SAVs compared to traditional transport modes such as conventional cars, urban PT (bus or tram), and bicycles. Second, we clarify whether the preferences of participants who could concentrate on both activities during the test rides differ from those who could focus on only one activity or none. Third, we describe how the choice between work and leisure activity influenced preferences for SAVs.

We organise the discussion of results according to the components of the model (Eq. (4)), starting with instrumental variables, then moving on to latent variables, and finally addressing socio-economic variables. Finally, using the estimated parameters for in-vehicle travel time and costs, we compute the corresponding VoTTs.

5.5.1. *Instrumental variables*

The study's findings from the main sample of participants reveal several patterns regarding the disutility of travel time and cost across four transportation alternatives (SAVs, conventional cars, urban PT (bus or tram), and bicycles), as well as the impact of engaging in work- and leisure-related NDRTs.

At first glance, the disutility of travel time for SAVs is similar to that of PT (bus or tram) and conventional cars, with participants' engagement in activities during the test rides showing no significant impact. This lack of difference may be attributed to the relatively small size of Delft, where travel distances and times are generally short.

However, when considering the ability to perform work or leisure activities while travelling in SAVs, participants show a clear preference for work-related tasks, which reduces the perceived disutility of travel time. Work tasks are seen as productive and purposeful, making travel time feel more valuable. In contrast, leisure activities may feel less rewarding. This distinction became apparent to participants only after experiencing both work and leisure activities firsthand during the test rides.

These findings align with prior research. Correia et al. (2019) highlighted a general preference for work over leisure activities in AVs, while Susilo et al. (2012) observed that performing work activities positively influenced train commuters' perceptions of travel time.

Additionally, those able to concentrate on both work and leisure during the rides perceived travel time in a conventional car more negatively. This may be explained by the fact that, unlike in SAVs, conventional car drivers must focus on driving and cannot make use of their travel time for other activities.

Table 10
Results of final joint mixed logit model with panel effects.

Parameter	Corresponding variable	Parameters Specific for pre-test	Specific for post-test	Generic
βik	**Instrumental variables**			
β_SAV_TT	In-vehicle travel time (min)	–	–	−0.0911***
β_PT_TT		–	–	−0.103***
β_CAR_TT		–	–	−0.0952***
β_BIKE_TT		–	–	−0.152***
Interaction effects with Activity during trip (Ref. − Leisure)				
β_SAV_TT * activity	Work	–	0.0139**	–
Interaction effects with Concentration on NDRTs (Ref. – Partial concentration or No concentration)				
β_CAR_TT * concentration	Full concentration	–	−0.0314***	–
β_SAV_TC	Travel costs (€)	–	–	−0.321***
β_CAR_TC		–	–	−0.439***
β_PT_TC		–	–	−0.527***
Crowdedness in PT (Ref. − Not crowded or Light crowdedness)				
β_PT_crowded	Crowded	−0.522*	–	–
βil	**Latent variables**			
β_PT_safety	Perceived safety, trust and enjoyment of AVs	–	1.15***	–
σ_safety	Standard deviation of error term	–	−0.000595	–
β_SAV_work	Perceived work activity	–	1.11**	–
β_PT_work		1.02***	–	–
σ_work	Standard deviation of error term	−0.155	0.00385	–
β_SAV_leisure	Perceived leisure activity	–	1.4**	–
β_CAR_leisure		–	1.42***	–
σ_leisure	Standard deviation of error term	–	0.853**	–
β_SAV_int	Intention to use SAVs	–	−1.47***	–
β_PT_int		−1.54***	–	–
σ_int	Standard deviation of error term	−0.398	0.0192	–
βis	**Socio-economic variables**			
Age (Ref. − Young)				
β_SAV_age	Old (above 50)	–	2.1***	–
Interaction effects with Belief in experimental set-up (Ref. – Hesitant or Did not believe)				
β_SAV_age * belief	Believed	–	−1.14***	–
β_BIKE_age * belief		–	0.766***	–
Education level (Ref. – Low)				
β_CAR_education	High	–	1.38***	–
Interaction effects with Belief in experimental set-up (Ref. – Hesitant or Did not believe)				
β_CAR_education * belief	Believed	–	−1.22***	–
Occupation (Ref. − Other)				
β_CAR_occupation	Employed	−0.947***	–	–
β_BIKE_occupation		–	1.04***	–
Interaction effects with Belief in experimental set-up (Ref. – Hesitant or Did not believe)				
β_BIKE_occupation * belief	Believed	–	−1.09***	–
Main transport mode (Ref. – Car or Other)				
β_PT_tr_mode_PT	PT	0.747**	–	–
β_BIKE_tr_mode_PT		−0.792**	–	–
Interaction effects with Concentration on NDRTs (Ref. – Partial concentration or No concentration)				
β_BIKE_tr_mode_PT * concentration	Full concentration	–	−1.59***	–
Interaction effects with Belief in experimental set-up (Ref. – Hesitant or Did not believe)				
β_CAR_tr_mode_PT * belief	Believed	–	−1.29***	–
Main transport mode (Ref. – PT or Other)				
β_CAR_tr_mode_car	Car	–	–	0.945***
β_PT_tr_mode_car		–	–	−1.26***
β_BIKE_tr_mode_car		−1.37***	−0.907***	–
Interaction effects with Concentration on NDRTs (Ref. – Partial concentration or No concentration)				
β_CAR_tr_mode_car * concentration	Full concentration	–	0.718**	–
Previous ride experience in AVs (Ref. − No)				
β_PT_ride_exp	Yes	–	2.83***	–
Interaction effects with Concentration on NDRTs (Ref. – Partial concentration or No concentration)				
β_PT_ride_exp * concentration	Full concentration	–	−2.09***	–
Interaction effects with Belief in experimental set-up (Ref. – Hesitant or Did not believe)				
β_PT_ride_exp * belief	Believed	–	−1.77**	–
Use of car- or ride-sharing services (Ref. − No)				
β_SAV_car_shar	Yes	0.766***	–	–
β_CAR_car_shar		0.688**	–	–
Motion sickness (Ref. − No)				
β_CAR_car_sick[a]	Yes	–	–	0.644***
Mobility restrictions (Ref. − No)				
β_SAV_mob_restr	Yes	−0.895***	–	–

(*continued on next page*)

Table 10 (*continued*)

Parameter	Corresponding variable	Parameters Specific for pre-test	Specific for post-test	Generic
Traffic accident (Ref. – No)				
β_CAR_tr_accident	Yes	−0.66***	–	–
μ	**Scale between waves**			
μ_wave2	Scale parameter	–	1.23*** [a]	–
σ	**Panel effects**			
σ	St. deviation for panel effects	0.684*	0.55	–
Number of parameters			51	
Sample size / Number of observations			104 / 1872	
Initial log-likelihood / Final log-likelihood			−2595.143 / −1800.632	
Rho-square / Adjusted Rho-square			0.306 / 0.287	
Akaike / Bayesian Information Criterion			3703.265 / 3985.538	
Number of Halton draws from a normal distribution			1000	

*** significant at a 99% confidence interval; ** at 95%; * at 90%; [a] *t*-test against 1.

[a] Generic variable between Motion sickness (before test rides) and Motion sickness (after test ride with leisure activity).

Table 11
Staged model estimation and goodness-of-fit statistics.

	Joint MNL model (instrumental variables)	Joint MNL model (instrumental and socio-economic variables)	Final joint ML model (instrumental, socio-economic, and latent variables)
Sample size		104	
Number of observations		1872	
Number of parameters	16	40	51
Initial log-likelihood		−2595.143	
Final log-likelihood	−2024.611	−1816.269	−1800.632
Rho-square	0.22	0.3	0.306
Adjusted Rho-square	0.214	0.285	0.287
Akaike Information Criterion	4081.222	3712.537	3703.265
Bayesian Information Criterion	4169.779	3933.928	3985.538

In contrast to the other three alternatives, the disutility of travel time for bicycles is significantly higher compared to other modes. This may be because cycling involves physical effort and lower comfort levels, so longer trips feel more tiring and less pleasant than travel by motorised modes.

Regarding travel costs, the parameters for SAVs, conventional cars and PT (bus or tram) remained stable before and after the test rides. Perceived travel costs were least negative for SAVs, followed by conventional cars, and most negative for PT. Similar patterns have been observed in earlier studies: PT is often associated with higher generalised costs due to fares, waiting, and transfers, making it less attractive compared to the private car (Wardman, 2004; Chowdhury et al., 2015). In contrast, emerging modes such as SAVs are perceived more favourably in cost terms, as they are expected to reduce or eliminate fixed expenses related to car ownership, parking, and maintenance (Nazari et al., 2018; Pakusch et al., 2018; Yap et al., 2016). This combination of lower fixed costs and higher convenience may explain why participants in this study viewed the cost of SAVs more positively than that of conventional cars and PT.

Additionally, the choice for the PT (bus or tram) alternative was negatively affected by the level of crowdedness before the test rides. Possibly, the novelty or positive aspects of the SAV experience could have reduced the weight of pre-existing concerns about crowdedness in public transport, making this PT attribute less influential in participants' choices.

Walking and waiting times were not significant factors in choosing between conventional car, PT (bus or tram), or SAV alternatives, possibly due to Delft's relatively small city size, which reduces the perceived burden of these factors.

5.5.2. Latent attitudinal variables

Latent attitudinal variables provide further insights into the travel mode preferences of participants based on underlying psychological factors. The four constructs, namely, perceived safety, trust and enjoyment of AVs, perception of work and leisure activities and intention to use SAVs, significantly influenced preferences for SAVs, conventional cars and urban PT (bus or tram), except for bicycles.

Before the test rides, participants' preference for PT, such as buses or trams, was positively influenced by the perception of work activities. This is likely because they were already familiar with working during their travel in these conventional modes. The relatively stable and predictable environment that PT offers, such as designated seating, consistent travel speeds, and the absence of driving responsibilities, may have made it an appealing option for those who value the ability to work during their journeys. After the test rides, preferences shifted toward SAVs based on the perception of work activities. This shift suggests that participants recognised SAVs could provide similar opportunities for productive use of travel time as PT, while also offering greater flexibility and privacy.

Based on the perception of leisure activities, participants favoured both SAVs and conventional cars after test rides without significantly different preferences between them, suggesting that cars, either conventional or automated, are viable for leisure on the way, possibly due to comfort and privacy. Unlike work activities, leisure may not require a stable environment, making both SAVs and traditional cars suitable choices for passengers who prioritise relaxation or entertainment during their journeys.

Moreover, participants who lack trust in AVs, perceive them as unsafe and generally have a negative opinion about them, tend to favour the PT transport alternative after the test rides. This preference for PT can be explained by the fact that public transport is more familiar and perceived as a safer option for those sceptical about AV technology.

Participants who initially intended to use SAVs in the future, also believing they could save them travel time, tend to avoid choosing PT. However, after gaining ride experience, their intentions shifted, leading them to no longer prefer SAVs as a mode of transport. This change could be due to the realisation that the expected time savings were not as substantial as they had thought, or they found other aspects of the SAV experience, such as comfort or convenience, less appealing than expected. This finding underscores the importance of experience in shaping perceptions and preferences, as perceived benefits may not always align with real-world experiences.

5.5.3. Socio-economic variables

Socio-economic factors also played a role in the preferences for SAVs, PT (bus or tram), conventional cars and bicycles, albeit in a limited way. Older participants, particularly those over 50, showed a preference for SAVs after test rides, although this preference was less pronounced among those who believed in the experimental setup. In addition, a significant interaction effect was found: older participants (over 50) who expressed belief in the experimental setup showed a stronger preference for the bicycle alternative. After the test rides, participants with higher education levels showed a positive preference for the car alternative, although this preference was reduced among those who believed in the experimental setup.

Employment status influenced choices, with employed participants initially showing a significant negative preference for cars, but shifting to a positive preference for bicycles after the test rides. Meanwhile, the use of car-sharing services predicted initial preferences for both SAVs and cars but lost significance after the test rides, indicating that direct experience with SAVs might diminish the appeal of shared services.

Predictably, current PT users preferred PT and avoided bicycles before the test rides. After the rides, those PT users who could fully concentrate on activities in SAVs continued to avoid bicycles, while those who believed in the experimental setup showed less preference for cars. In contrast, habitual car users consistently preferred their own mode and avoided PT (no significant difference between pre- and post-test rides). Their aversion to bicycles was stronger before the test rides but became less negative afterwards, and the preference for cars was further reinforced among those able to fully concentrate on activities during the SAV rides. These findings suggest that while direct experiences can shift perceptions and preferences to some extent, deeply ingrained habits and personal circumstances still significantly influence travel choices.

Additionally, participants with mobility restrictions tended not to prefer SAVs before the test rides, possibly due to concerns about accessibility in an automated driving environment. Similarly, individuals who had previously been involved in traffic accidents demonstrated an aversion to cars before the test rides, likely reflecting heightened safety concerns or anxiety associated with car travel, concerns that lost their significance after the test rides. Participants who reported a tendency to experience motion sickness before test rides and after the test ride with leisure activity showed a significantly higher likelihood of choosing the car alternative. Interestingly, participants with prior ride experience in AVs showed a preference for PT after the test rides. However, this preference was weaker among those who could fully concentrate on activities during the rides and among those who believed in the experimental setup. Overall, these results underscore how past experiences and personal circumstances influence how people evaluate and choose between emerging mobility options like SAVs.

5.5.4. Value of travel time

Using the disutilities of travel time and costs, we calculated the corresponding VoTTs for the SAV, PT (bus or tram) and conventional car alternatives, except for bicycles, where we assumed no associated travel costs (Table 12). When compared to the average VoTTs in the Dutch population (Kouwenhoven et al., 2023), which are 10.42 €/h for cars and 7.12 €/h for PT, the pattern of higher VoTTs for cars (13.0 €/h) and lower ones for PT (11.7 €/h) is consistent, albeit at a higher scale. This higher scale is most naturally a result of a very specific population, the one of Delft, being part of the experiment and not the average population in the Netherlands.

Examining further the impact of using travel time for NDRTs in SAVs, we observe the anticipated reduction of the VoTT for the SAV

Table 12
Value of travel time before and after test rides.

Transport mode	Travel time (min)		Travel costs (€)		Value of travel time (€/h)	
	Wave 1. Before test rides	Wave 2. After tests rides	Wave 1. Before test rides	Wave 2. After tests rides	Wave 1. Before test rides	Wave 2. After tests rides
Shared automated vehicle	−0.0911		−0.321		17.0	
	Travel time x Work activity					
	Not sign.	0.0139			–	14.4
Public transport (bus or tram)	−0.103		−0.527		11.7	
Conventional car	−0.0952		−0.439		13.0	
	Travel time x Full concentration					
	n/a	−0.0314				17.3

option (14.4 €/h) when travellers can engage in work-related activities (interaction variable). This finding aligns with earlier observations that VoTTs in traditional PT decrease when travel time is utilised productively for work (Kouwenhoven & de Jong, 2018; Molin et al., 2020). Importantly, this interaction effect only became significant in the post-test, after participants experienced the experimental SAV rides, underlying the role of exposure in revealing realistic trade-offs. Conversely, for those participants who could focus on both work and leisure activities (interaction variable with Full concentration) during their experimental SAV test rides, the VoTT for the conventional car alternative increased. This increase among some participants could be due to their heightened expectations of travel comfort and productivity, set by their positive experiences with multitasking in the experimental SAVs.

5.6. Comparison between the main sample and control group

When comparing the main sample of participants (pre-test responses) to the control group, several notable similarities and differences emerge in their travel preferences for the SAV, PT (bus or tram), conventional car and bicycle alternatives (Table 13). Both groups exhibit similar perceptions of in-vehicle travel time for all four travel options and a preference for current modes of transport, with habitual car users typically opting for cars and expressing a dislike for PT. Notably, factors such as activity on the way (work or leisure), waiting and walking times did not significantly influence mode choice across different transport alternatives in either sample.

Differences between the groups, however, highlight distinct influences on transport preferences. In the control group, travel costs did not significantly affect preferences for cars, PT (bus or tram), or SAVs, contrasting with the main sample. Crowded environments significantly deterred only participants from the main sample from choosing PT.

Demographic factors also played a role: female participants in the control group showed a preference for bicycles, while older participants avoided this transport alternative, unlike in the main sample, where older participants tended to shun PT. Higher education levels in the control group correlated with a preference for bicycles over cars, and higher income was a deterrent for selecting PT. While the use of shared services was a positive factor in choosing SAVs and cars in the main sample, it negatively affected SAV choices in the control group, who instead showed a preference for bicycles. Participants with motion sickness preferred cars in the main sample but bicycles in the control group. Unique to the main sample, mobility restrictions negatively impacted the choice of SAVs, possession of a driving license discouraged PT use, and past traffic accidents deterred car use.

In conclusion, the comparison between the main sample (pre-test responses) and the control group revealed that both groups shared common perceptions of in-vehicle travel time and mode loyalty, and were similarly unaffected by factors such as activity on the way (work or leisure), waiting times, and walking times. However, the main sample showed greater sensitivity to factors like travel costs and crowding. Socioeconomic characteristics, on the other hand, did not yield consistent patterns across the two groups. Longitudinal studies conducted during the first application cases of SAVs could help identify the socio-economic profiles of potential users more clearly.

5.7. Study limitations and future research directions

This study faces several limitations that should be considered when interpreting the results.

Two limitations stem from the SC experiment. First, participants were asked to make decisions based on hypothetical scenarios, which becomes challenging when envisioning innovative transport options like SAVs, with which they have no prior experience. Second, the study focused exclusively on SAVs operating as car-sharing services, excluding ride-sharing or pooling scenarios, which remain less preferred due to barriers like the discomfort of sharing confined spaces with strangers (Correia & Viegas, 2011; Hamadneh & Esztergár-Kiss, 2022).

To mitigate the challenge of limited familiarity with SAVs, participants were provided with a ride experience in a Wizard-of-Oz simulator-on-wheels vehicle, a proven tool for simulating AVs (Detjen et al., 2020; Dillmann et al., 2023; Klingegård et al., 2020). However, the experimental setting itself introduced limitations. The presence of a 'safety driver' and the vehicle's setup (partitions, screens, and cameras) led to suspicions among 26% of participants. Additionally, variability in driving styles from seven test drivers may have influenced ride experiences. To reduce this effect, drivers were instructed to maintain a calm and cautious style to mimic AV behaviour to minimise the variability. However, it is important to note that evaluating driving style was not the objective of this study. Instead, the focus was on passenger perceptions and behavioural responses under the belief that the vehicle was operating autonomously. Future studies using real SAVs are therefore needed to validate the findings and assess how genuine automated driving behaviour shapes user perceptions and preferences.

As shown in Table B.3, the vast majority of participants had no prior experience with automated driving systems. This reflects the current European context, where SAVs are not yet publicly available on open roads. As such, the reactions captured in this study are likely representative of how early-stage users might respond when first experiencing SAVs.

All rides were conducted during daylight hours, meaning the effect of low-light or nighttime conditions on NDRT performance was not evaluated. Future research should explore how lighting conditions, particularly the absence of dedicated interior lighting in SAVs, may affect task engagement and user preferences, especially during early morning or evening travel.

It should be noted that the experiment was conducted in late 2022, when residual pandemic-related behavioural adaptations, such as heightened sensitivity to shared spaces, changes in public transport use, and increased familiarity with teleworking, may have influenced participants' baseline attitudes toward SAVs. Future studies could replicate this experiment in post-pandemic conditions to disentangle experience effects from residual COVID-related travel attitudes.

The study's moderate sample size presents another limitation, stemming from the effort to validate user attitudes and preferences for SAVs in a field experiment while operating under time and budget constraints. Furthermore, self-selection bias may have influenced

Table 13
Comparison between the main sample and control group: model results.

Parameter	Corresponding variable	Parameters Specific for pre-test	Specific for control group	Generic
βik	**Instrumental variables**			
β_SAV_TT	In-vehicle travel time (min)	–	–	−0.0617***
β_CAR_TT		–	–	−0.11***
β_PT_TT		–	–	−0.0991***
β_BIKE_TT		–	–	−0.138***
β_SAV_TC	Travel costs (€)	−0.358***	–	–
β_CAR_TC		−0.616***	–	–
β_PT_TC		−0.559***	–	–
Crowdedness in PT (Ref. – Not crowded or Light crowdedness)				
β_PT_crowded	Crowded	−0.483*	–	–
βis	**Socio-economic variables**			
Gender (Ref. – Male)				
β_BIKE_gender	Female	–	0.484*	–
Age (Ref. – Young)				
β_PT_age	Old (above 50)	−0.567**	–	–
β_BIKE_age		–	−1.0***	–
Education level (Ref. – Low)				
β_CAR_edu	High	–	−1.28***	–
β_BIKE_edu		–	1.4***	–
Occupation (Ref. – Other)				
β_CAR_occupation	Employed	−1.01***	–	–
Income (gross annual per household) (Ref. – Blow 50 k)				
β_PT_income	Above 50 k	–	−0.907**	–
Main transport mode (Ref. – Car or Other)				
β_PT_tr_mode_PT	PT	0.697**	–	–
β_BIKE_tr_mode_PT		−0.67**	–	–
Main transport mode (Ref. – PT or Other)				
β_CAR_tr_mode_car	Car	–	–	1.37***
β_PT_tr_mode_car		–	–	−0.838***
β_BIKE_tr_mode_car		−1.06***	–	–
Use of car- or ride-sharing services (Ref. – No)				
β_SAV_car_shar	Yes	0.638***	−0.517*	–
β_CAR_car_shar		0.706**	–	–
β_BIKE_car_shar		–	0.821**	–
Motion sickness (Ref. – No)				
β_CAR_car_sick	Yes	0.621*	–	–
β_BIKE_car_sick		–	1.35***	–
Mobility restrictions (Ref. – No)				
β_SAV_mob_restr	Yes	−1.5***	–	–
Driving license (Ref. – No)				
β_PT_dr_license	Yes	−0.96***	–	–
Traffic accident (Ref. – No)				
β_CAR_tr_accident	Yes	−0.683***	–	–
μ	**Scale between waves**			
μ_wave0	Scale parameter	–	1.14*** [a]	–
σ	**Panel effects**			
σ_μ	Standard deviation for panel effects	1.28***	0.951	–
Number of parameters			32	
Sample size main group / Number of observations			104 / 936	
Sample size control group / Number of observations			35 / 315	
Initial log-likelihood / Final log-likelihood			−1734.254 / −1248.155	
Rho-square / Adjusted Rho-square			0.28 / 0.262	
Akaike / Bayesian Information CriterionNumber of Halton draws from a normal distribution			2560.31 / 2724.5241000	

*** significant at a 99% confidence interval; ** at 95%; * at 90%; [a] *t*-test against 1.

the participant pool, as members of the Delft panel registered within three hours after the study advertisement was released, suggesting that those curious or enthusiastic about AVs were more inclined to participate. These individuals, presumably early adopters, may not fully represent the general population (Feys et al., 2021). To address this, a control group of participants without the incentive to ride in an AV was included.

Participants' awareness of being observed during the test rides could cause the Hawthorne effect, potentially altering their behaviour (Skippon et al., 2016). Finally, since this study provided only a short-term experience, participants' attitudes and preferences may change over time. This underscores the importance of conducting longitudinal studies in real-world application cases and provides a natural transition to sharing potential directions for future research.

Future studies could explore the integration of SAVs into diverse built environments, including rural areas and cities of varying sizes, to assess differences in user attitudes and preferences. Additionally, researchers could examine how the use of travel time for

NDRTs differs depending on whether SAVs are used as a primary transport mode or as part of a multimodal journey, particularly for access and egress to transit lines. Understanding which existing travel modes users may switch from when adopting SAVs, as explored by Öztürker et al. (2022), is another crucial area of investigation.

Furthermore, future research should also consider passive activities, such as sleeping or simply observing the surroundings. The extent to which users will choose to spend travel time productively in SAVs remains an open question (Pudāne et al., 2019; Singleton, 2019). Future studies could also extend the assessment of NDRT engagement levels by combining non-intrusive observation of self-selected everyday activities (e.g., reading, emailing) captured by video recordings with more direct validation measures, such as brief post-ride comprehension checks or standardised tasks during the test rides (e.g., n-Back) and gaze-based indicators like eye-tracking or blink rate.

Future research could also explore more advanced model specifications to better capture heterogeneity in unobserved utility variance. In particular, random scale mixed logit models or latent class models (Hess & Train, 2017) allow for more flexible treatment of scale by accounting for individual- or subgroup-level variation in error structures. While such approaches introduce greater model complexity and require larger sample sizes, they offer valuable opportunities to uncover latent heterogeneity that may be masked in fixed-scale formulations. These extensions would be especially relevant for studies aiming to segment user preferences or test responses to SAV exposure across diverse population groups.

Addressing these limitations and pursuing the suggested research directions will provide a deeper understanding of user preferences, behavioural shifts, and the long-term impact of travel time use in SAVs across various environments and use cases, ultimately aiding in the successful integration of AVs into future public transport systems.

6. Conclusions

Shared automated vehicles (SAVs) operating in urban areas have the potential to transform travel by enabling users to engage in non-driving-related tasks (NDRTs), enhancing productivity and travel satisfaction. When integrated with public transport systems, SAVs can offer a more sustainable alternative to privately owned AVs, while maintaining similar levels of privacy and comfort in a less crowded environment.

In the present study, we explored how the possibility of using travel time for work and leisure NDRTs in SAVs within urban environments influences users' attitudes and preferences for these vehicles when compared to traditional travel modes such as conventional cars, urban PT (buses or trams), and bicycles. We also looked at the important economic indicator of the perceived value of travel time (VoTT). To understand users' attitudes and preferences for an SAV service, we conducted a field study using a Wizard-of-Oz (WoZ) simulator-on-wheels vehicle to emulate such a service being available on the streets of Delft (Netherlands). A total of 104 participants completed two test rides, engaging in work- and leisure-related NDRTs while their engagement levels were measured via video recordings.

Results of a mixed panel logit model estimated jointly on pre- and post-test data revealed several key findings that deepen the understanding of behavioural and experiential aspects of travel time use in SAVs. SAVs were perceived as less sensitive to travel costs than PT and cars, while preference for work-related NDRTs reduced the negative perception of travel time in SAVs post-test, aligning with prior findings that productive use of travel time enhances travel satisfaction. Moreover, engagement levels during SAV rides further influenced preferences, as participants who fully concentrated on both work and leisure activities perceived travel time in conventional cars more negatively. These results reflect how direct exposure to work and leisure NDRTs reshapes preferences across modes.

The shifts in the perceptions of travel time were reflected in changes to VoTTs. Before the test rides, SAVs had higher VoTTs than cars and PT. After the rides, VoTTs for SAVs decreased when used for work-related activities, underscoring their advantage for productivity-focused travel. By contrast, the ability to fully concentrate on NDRTs during test rides raised VoTTs for conventional cars to match SAVs' levels for leisure-oriented participants. PT VoTTs remained unchanged.

The results also revealed changes in participants' attitudes toward SAVs after their firsthand experiences. Initial participants' preferences shifted from the PT pre-test to the SAVs post-test, driven by the perception of using travel time for work activity, which includes the comfort of conducting work tasks and the ability to concentrate on them. While PT was initially preferred due to participants' familiarity with working during commutes, the experience of NDRTs during the test rides enhanced their perception of SAVs as a productive and comfortable environment for work.

Based on the perception of using travel time for leisure activities, both SAVs and conventional cars were equally favoured post-test, highlighting their comfort and privacy for more relaxing activities. For those sceptical of AV technology, trust and safety concerns led to a preference for PT as a more familiar and safer alternative. Additionally, participants who initially intended to use SAVs for time savings tended to avoid PT, but after the rides, no longer showed interest in SAVs as their preferred mode of transport.

Lastly, the WoZ experiment proved to be a relevant method for simulating realistic SAV experiences, with 74% of participants believing they were riding in an AV. This approach provided participants with the opportunity to experience work- and leisure-related NDRTs while travelling, offering a practical framework for future behavioural research in urban contexts.

In summary, this study demonstrates how an immersive WoZ experimental framework can advance our understanding of travel behaviour in SAVs. Compared to hypothetical stated-preference surveys, this exposure captured behavioural responses that aligned more closely with revealed-preference evidence (Kouwenhoven & de Jong, 2018; Molin et al., 2020) and showed shifts in preferences, attitudes, and VoTTs tied to real NDRT engagement. By capturing these experiential factors, the study provides a nuanced perspective on how user expectations are formed and highlights SAVs' potential to redefine productive travel time use while paving the way for their integration into urban transport systems.

Beyond immediate impacts on user attitudes and preferences, the findings underscore the broader societal and environmental implications of SAV adoption. The ability of SAVs to promote productive use of travel time could lead to shifts in urban mobility patterns, encouraging fewer private car trips and enhancing the efficiency of shared mobility systems.

Although the findings of this study provide preliminary evidence based on short-term experiences with SAVs in experimental settings, they highlight the importance of accounting for the use of travel time in this travel mode to accurately capture user attitudes and preferences. This consideration is particularly relevant for public transport operators, urban policymakers, and private mobility providers when making investment decisions aimed at improving travel comfort and facilitating productive use of travel time.

CRediT authorship contribution statement

Maryna Öztürker: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sina Nordhoff:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Sascha Hoogendoorn-Lanser:** Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Bart van Arem:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Gonçalo Homem de Almeida Correia:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Given his role as Associate Editor, Gonçalo Correia had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The experiment was financially supported by hEAT Lab and Transport Institute (Delft University of Technology). The authors are grateful to LeasePlan for providing the vehicle, the municipality of Delft for their help with the distribution of the call for the experiment, cafe Labs for hosting our team and participants, Koen de Clercq for help with the translation of the surveys to Dutch, Peter van Oossanen and Edwin Scharp for technical support and installation of equipment, Nadia Haifi for catchy advertisements and the team of drivers (Jim Hoogenboom, Mathijs Stultjens, Frerik Zandhuis, Rody Boting, Kevin Duijn, Jens van den Berg, Nicola Manservisi). Special gratitude goes to our participants who dared to participate.

Appendix A

Example NDRTs provided to participants

Work-related tasks

- Reading and replying to emails
- Reviewing documents or reports
- Scheduling appointments or planning a weekly agenda
- Writing a letter, memo, or text document
- Reading work materials
- Preparing a to-do list or grocery list

Leisure-related tasks

- Reading books, magazines, or news articles
- Watching videos or listening to music/podcasts
- Playing mobile games
- Browsing social media or messaging friends

Participants were free to bring their own materials (e.g., laptop, smartphone, paper documents) or choose from a selection of books and devices provided on-site.

Appendix B

Table B1
Overview of the sample and control group and comparison with the distribution in the Dutch population.

Variable	Category (model)	Category (survey)	Sample, % (104 respondents)	Control group, % (35 respondents)	Population, % (CBS, 2021; CBS 2022a,b; CBS, 2023a,b)
Gender	Male (reference)	Male	65.4	62.9	49.7
	Female	Female	34.6	37.1	50.3
		Prefer not to say	–	–	–
Age	Young (reference)	18–29	12.5	14.2	15.4
		30–39	4.8	34.6	12.6
		40–49	10.6	34.6	12.1
	Old	50–59	21.2	8.6	15.9
		60–69	30.8	–	12.4
		≥70	20.2	–	31.6
		Prefer not to say	–	8.6	–
Education level	Low (reference)	No education	–	–	–
		Primary education	1.0	–	9.0
		Secondary education	16.3	2.9	29.6
		Higher National Diploma	9.6	2.9	26.6
	High	Undergraduate degree (Bachelor's degree)	29.8	17.1	21.2
		Postgraduate degree or higher (Master's degree, PhD, etc.)	43.3	74.2	13.6
		Prefer not to say	–	2.9	–
Income (gross annual per household)	Below 50 k (reference)	Less than €9,999	1.9	8.7	4.1
		€10,000 – €19,999	4.8	5.7	11.3
		€20,000 – €29,999	6.7	–	31.5
		€30,000 – €39,999	8.7	11.4	26.1
		€40,000 – €49,999	4.8	–	14.5
	Above 50 k	€50,000 – €59,999	16.3	5.7	6.3
		€60,000 – €69,999	5.8	5.7	2.7
		≥ € 70,000	34.6	51.5	3.5
		Prefer not to say	16.3	11.4	–
Occupation	Employed	Employed or self-employed	50.0	94.2	74.2
	Others (reference)	Unemployed or (partially) incapacitated	3.8	–	2.7
		Retired	28.8	–	11.8
		Student or intern	4.8	2.9	2.5
		Housewife/houseman	1.0	2.9	2.1
		Volunteer	6.7	–	6.7 (2 last categories)
		Others	4.8	–	

Table B2
Distribution of main transport modes according to the trip purpose in the sample, control group and the Dutch population.

Transport mode / Trip purpose*	Category (model)	Work, %	Study, %	Frequently visited destination, %	Sample (control group)[1], %	Population, % (CBS, 2022c)
Car (as a driver)	Car	16.3 (37.1)	– (–)	22.2 (2.9)	38.5 (40.0)	50.0
Car (as a passenger)		1.0 (2.9)	– (–)	0.9 (–)	1.9 (2.9)	18.1
Train	PT	5.8 (2.9)	1.0 (–)	2.8 (–)	9.6 (2.9)	8.7
Bus, tram, or metro		1.0 (8.5)	– (–)	2.8 (–)	3.8 (8.5)	2.6
Bicycle or e-bike	Others (reference)	21.2 (40.0)	2.8 (–)	12.5 (–)	36.5 (40.0)	10.1
Scooter		1.0 (–)	– (–)	– (–)	1.0 (–)	n/a [4]
Walking		1.9 (–)	1.0 (–)	3.8 (–)	6.7 (–)	4.0
Working from home [2]		1.9 (2.9)	– (2.9)	– (–)	1.9 (5.7)	n/a [4]
Other [3]		–	–	–	–	6.5

[1] The data for the control group are given in parentheses.
[2] Frequently visited destination was used as the purpose of the trip for this category of respondents.
[3] This category is absent from the survey.
[4] These categories are missing in the statistics.
* The trip purpose of participants is used as the context in the SC experiment.

Table B3
Additional characteristics of the participants in the main experiment and the control group.

Variable	Category (model)	Category (survey)	Sample, %	Control group, %
Driving license	Yes	Yes	91.3	91.4
	No (reference)	No	8.7	2.9
		Prefer not to say	–	5.7
PT pass	Yes	Yes	92.3	85.7
	No (reference)	No	7.7	11.4
		Prefer not to say	–	2.9
Traffic accident	Yes	Yes	60.5	51.4
	No (reference)	No	38.5	45.7
		Prefer not to say	1.0	2.9
Use of car- or ride-sharing services (such as Uber and Green Wheels)	Yes	Yes, regularly	1.0	2.9
		Yes, occasionally	17.3	71.3
	No (reference)	No, never used it	76.9	22.9
		No, not familiar with these services	4.8	–
		Prefer not to say	–	2.9
Motion sickness (before test rides)	Yes	Yes	18.3	25.7
	No (reference)	No	81.7	74.3
		Prefer not to say	–	–
Motion sickness (during test rides with work activity)	Yes	Extremely	1.9	n/a
		Moderately	4.8	
		Somewhat	6.7	
		Slightly	12.5	
	No (reference)	Not at all	74.1	
Motion sickness (during test rides with leisure activity)	Yes	Extremely	2.9	n/a
		Moderately	3.8	
		Somewhat	5.8	
		Slightly	18.3	
	No (reference)	Not at all	69.2	
Mobility restrictions	Yes	My mobility is limited (due to age, pregnancy, etc.)	3.8	–
		I am not mobile (due to health problems, age, etc.)	1.0	–
	No (reference)	I am fully mobile (physically)	95.2	100.0
		Prefer not to say	–	–
Total travel time	Long travel	More than 30 min	31.7	17.2
	Short travel	15–30 min	26.9	25.7
	(reference)	Less than 15 min	40.4	57.1
Frequency of travel	Frequent	4 or more days per week	40.4	57.1
	Others (reference)	1–3 days per week	45.2	34.3
		1–3 days per month	9.6	–
		1–3 days in the last three months	2.9	–
		I work from home	1.9	8.6
Ride experience in AVs	Yes	Yes, more than once	2.9	2.9
		Yes, once	6.7	8.6
	No (reference)	No, never	90.4	88.5

Appendix C

Table C1
Results of confirmatory factor analysis: model development stages (pre-test).

Stages of model development	First run of the model (hypothesised structure)	Second run	Third run	Final model
Factor 1. Enjoyment of AVs			**Factors 1 and 2 combined**	
S1	0.697	0.749	0.663	0.648
S2	0.633***	0.627***	0.547***	
S3	0.322***			
Factor 2. Perceived safety and trust				
S4	0.853	0.868	0.857***	0.873***
S5	0.492***			
S6	0.802***	0.809***	0.791***	0.798***
S7	0.833***	0.839***	0.836***	0.829***
Factor 3. Work activity				
S8	0.507			
S9	0.685***	0.692	0.702	0.700
S10	0.902***	0.923***	0.912***	0.911***
S11	0.375***			
Factor 4. Leisure activity				
S8	0.532			
S12	0.645***	0.626	0.625	0.620
S13	0.721***	0.713***	0.724***	0.726***
S11	0.807***	0.814***	0.803***	0.801***
Factor 5. Intention to use shared automated vehicles				
S14	0.512			
S15	0.853***	0.841	0.825	0.819
S16	0.622***	0.602***	0.605***	0.610***
Factor 6. Service quality				
S17	0.359	–	–	–
S18	0.967 (p = 0.125)	–	–	–
Model fit	RMSEA = 0.078	RMSEA = 0.054	RMSEA = 0.086	RMSEA = 0.081
	CFI = 0.97	CFI = 0.983	CFI = 0.952	CFI = 0.965
	TLI = 0.951	TLI = 0.971	TLI = 0.927	TLI = 0.942

Table C2
Assessment of discriminant validity: second-stage CFA model (pre-test).

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 1	**0.693**	0.853	0.573	0.658	0.605
Factor 2	0.853	**0.839**	0.681	0.624	0.428
Factor 3	0.573	0.681	**0.819**	0.571	0.357
Factor 4	0.658	0.624	0.571	**0.722**	0.676
Factor 5	0.605	0.428	0.357	0.676	**0.831**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Table C3
Assessment of discriminant validity: final CFA model (pre-test).

	Factor 1 (combined Factors 1 and 2)	**Factor 2** (previously Factor 3)	**Factor 3** (previously Factor 4)	**Factor 4** (previously Factor 5)
Factor 1	**0.794**	0.691	0.701	0.482
Factor 2	0.691	**0.812**	0.581	0.375
Factor 3	0.701	0.581	**0.721**	0.69
Factor 4	0.482	0.375	0.69	**0.721**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Table C4
Results of confirmatory factor analysis: model development stages (post-test).

Stages of model development	First run of the model (hypothesised structure)	Second run	Third run	Final model
Factor 1. Enjoyment of AVs			**Factors 1 and 2 combined**	
S1	0.896	0.882	0.895	0.728
S2	0.647***	0.626***	0.591***	
S3	0.387***			
Factor 2. Perceived safety and trust				
S4	0.912	0.796	0.660***	0.768***
S5	0.549***			
S6	0.86***	0.689***	0.654***	0.672***
S7	0.71***	0.800***	0.722***	0.776***
Factor 3. Work activity				
S8	0.374			
S9	0.844***	0.847	0.836	0.851
S10	0.939***	0.939***	0.951***	0.932***
S11	0.415***			
Factor 4. Leisure activity				
S8	0.503			
S12	0.873***	0.648	0.609	0.664
S13	0.916***	0.722***	0.716***	0.735***
S11	0.702***	0.769***	0.776***	0.755***
Factor 5. Intention to use shared automated vehicles				
S14	0.42			
S15	0.758***	0.743	0.770	0.740
S16	0.829***	0.839***	0.809***	0.842***
Factor 6. Service quality				
S17	0.391	–	–	–
S18	1.211 (p = 0.146)	–	–	–
Model fit	RMSEA = 0.097 CFI = 0.893 TLI = 0.862	RMSEA = 0.057 CFI = 0.985 TLI = 0.974	RMSEA = 0.096 CFI = 0.952 TLI = 0.927	RMSEA = 0.075 CFI = 0.972 TLI = 0.954

Table C5
Assessment of discriminant validity: second-stage CFA model (post-test).

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 1	**0.765**	0.859	0.451	0.552	0.654
Factor 2	0.859	**0.764**	0.382	0.609	0.708
Factor 3	0.451	0.382	**0.894**	0.665	0.522
Factor 4	0.552	0.609	0.665	**0.716**	0.603
Factor 5	0.654	0.708	0.522	0.603	**0.792**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Table C6
Assessment of discriminant validity: final CFA model (post-test).

	Factor 1 (combined Factors 1 and 2)	**Factor 2** (previously Factor 3)	**Factor 3** (previously Factor 4)	**Factor 4** (previously Factor 5)
Factor 1	**0.735**	0.44	0.638	0.761
Factor 2	0.44	**0.894**	0.658	0.524
Factor 3	0.638	0.658	**0.721**	0.608
Factor 4	0.761	0.524	0.608	**0.794**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Table C7
Results of confirmatory factor analysis: model development stages (control group).

Stages of model development	First run of the model (hypothesised structure)	Second run	Third run	Fourth run	Final model
Factor 1. Enjoyment of AVs				**Factors 1 and 2 combined**	
S1	0.62	0.685	0.701	0.717	0.696
S2	0.817***	0.879***	0.944***	0.908***	0.951***
S3	0.627***	0.536***			

(*continued on next page*)

Table C7 (*continued*)

Stages of model development	First run of the model (hypothesised structure)	Second run	Third run	Fourth run	Final model
Factor 2. Perceived safety and trust					
S4	0.853	0.889	0.907	0.627***	0.562***
S5	0.602***	0.632***	0.628***	0.46***	
S6	0.149 (p = 0.391)				
S7	0.933***	0.886***	0.868***	0.545***	
Factor 3. Work activity					
S8	0.055				
S9	1.975 (p = 0.493)				
S10	0.401 (p = 0.127)				
S11	0.074 (p = 0.162)				
Factor 4. Leisure activity					
S8	0.683	0.684	0.672	0.672	0.672
S12	0.705***	0.707***	0.71***	0.727***	0.709***
S13	0.964***	0.965***	0.968***	0.958***	0.969***
S11	0.703***	0.717***	0.72***	0.721***	0.719***
Factor 5. Intention to use shared automated vehicles					
S14	0.409				
S15	0.56***				
S16	0.556***				
Factor 6. Service quality					
S17	0.435				
S18	0.365***				
Model fit	RMSEA = 0.175 CFI = 0.761 TLI = 0.688	RMSEA = 0.133 CFI = 0.905 TLI = 0.867	RMSEA = 0.114 CFI = 0.943 TLI = 0.914	RMSEA = 0.124 CFI = 0.948 TLI = 0.917	RMSEA = 0.122 CFI = 0.953 TLI = 0.924

Table C8
Assessment of discriminant validity: third-stage CFA model (control group).

	Factor 1	**Factor 2**	**Factor 4**
Factor 1	**0.831**	0.926	0.59
Factor 2	0.926	**0.811**	0.544
Factor 4	0.59	0.544	**0.777**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Table C9
Assessment of discriminant validity: final CFA model (control group).

	Factor 1 (combined Factors 1 and 2)	**Factor 3** (previously Factor 4)
Factor 1	**0.755**	0.538
Factor 3	0.538	**0.775**

Diagonal values (bold) are the square root of the average variance extracted from the loadings.
Off-diagonal values are latent factor correlations.

Data availability

The anonymised survey data that underlie the analyses in this article can be obtained from the corresponding author upon request. Video recordings and raw interview transcripts are not publicly available due to ethical and confidentiality constraints agreed with participants at the time of data collection. Only derived variables (i.e., indicators of concentration on the task coded from video and belief in the experimental setup from interviews) are retained in the shared dataset.

References

Acheampong, R.A., Cugurullo, F., Gueriau, M., Dusparic, I., 2021. Can autonomous vehicles enable sustainable mobility in future cities? Insights and policy challenges from user preferences over different urban transport options. Cities 112, 103134. https://doi.org/10.1016/j.cities.2021.103134.
Bahamonde-Birke, F., de Dios Ortúzar, J., 2014. Is sequential estimation a suitable second best for estimation of hybrid choice models? Transp. Res. Rec. 2429 (1), 51–58. https://doi.org/10.3141/2429-06.
Bala, H., Anowar, S., Chng, S., Cheah, L., 2023. Review of studies on public acceptability and acceptance of shared autonomous mobility services: past, present and future. Transp. Rev. 1–27. https://doi.org/10.1080/01441647.2023.2188619.
Baltodano, S., Sibi, S., Martelaro, N., Gowda, N., Ju, W., 2015. The RRADS platform: a real road autonomous driving simulator. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 281–288. https://doi.org/10.1145/2799250.2799288.

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.
Barreto, R.A., Kellett, J., Pierce, J.C., Van den Hengel, A., 2022. I'm in love with my car psychological attachment to cars, automated vehicles and the driverless future. JTEP 56 (1), 28–55.
Bellem, H., Thiel, B., Schrauf, M., Krems, J.F., 2018. Comfort in automated driving: an analysis of preferences for different automated driving styles and their dependence on personality traits. Transport. Res. F: Traffic Psychol. Behav. 55, 90–100. https://doi.org/10.1016/j.trf.2018.02.036.
Bierlaire, M., 2023. A short introduction to Biogeme. Technical report TRANSP-OR 230620. Transport and Mobility Laboratory, ENAC, EPFL.
Britten, N., 2021. *An on-road assessment of driver secondary task engagement and performance during Assisted & Automated Driving* (Doctoral dissertation, Virginia Tech).
Brown, G.T., Harris, L.R., O'Quin, C., Lane, K.E., 2017. Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. Int. J. Res. Method Educ. 40 (1), 66–90. https://doi.org/10.1080/1743727X.2015.1070823.
Brown, T.A., 2015. Confirmatory Factor Analysis for Applied Research, 2nd ed. The Guilford Press, New York.
Carrese, F., Sportiello, S., Zhaksylykov, T., Colombaroni, C., Carrese, S., Papaveri, M., Patella, S.M., 2023. The integration of shared autonomous vehicles in public transportation services: a systematic review. Sustainability 15 (17), 13023. https://doi.org/10.3390/su151713023.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2021. Verdeling gestandaardiseerd inkomen [Income distribution (standardised income)]. Retrieved from https://www.cbs.nl/nl-nl/visualisaties/inkomensverdeling.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2022a. Bevolking; hoogstbehaald onderwijsniveau en onderwijsrichting, 2003-2022 [Population; highest level of education and direction of education achieved, 2003-2022]. Retrieved from https://www.cbs.nl/nl-nl/cijfers/detail/85184NED?dl=6B41D.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2022b. Bevolking op 1 januari en gemiddeld; geslacht, leeftijd en regio [Population on January 1 and average; gender, age and region]. Retrieved from https://www.cbs.nl/nl-nl/cijfers/detail/03759ned?dl=39E0B.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics] (2022c). Hoeveel wordt er met het openbaar vervoer gereisd? [How much travel is done by public transport?]. Retrieved from https://www.cbs.nl/nl-nl/visualisaties/verkeer-en-vervoer/personen/openbaar-vervoer.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2023a. Aantal mensen buiten beroepsbevolking in 2022 snel geslonken [Number of people outside the labor force will shrink rapidly in 2022]. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2023/18/aantal-mensen-buiten-beroepsbevolking-in-2022-snel-geslonken.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2023b. Bevolking, 15 tot 75 jaar. Samenstelling beroepsbevolking [Population, 15 to 75 years. Composition of the labour force]. Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-beroepsbevolking/bevolking-15-tot-75-jaar.
Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics], 2023c. Mobility; per trip, trip characteristics, modes of travel and regions. Retrieved from.
Chen, R., Arief, M., Zhang, W., Zhao, D., 2020. How to evaluate proving grounds for self-driving? a quantitative approach. IEEE Trans. Intell. Transp. Syst. 22 (9), 5737–5748. https://doi.org/10.1109/TITS.2020.2991757.
Cheung, G.W., Cooper-Thomas, H.D., Lau, R.S., Wang, L.C., 2024. Reporting reliability, convergent and discriminant validity with structural equation modeling: a review and best-practice recommendations. Asia Pac. J. Manag. 41 (2), 745–783. https://doi.org/10.1007/s10490-023-09871-y.
ChoiceMetrics, 2018. Ngene 1.2. User manual & reference guide. The cutting edge in experimental design. Retrieved from http://www.choice-metrics.com/NgeneManual120.pdf.
Chowdhury, S., Ceder, A.A., Schwalger, B., 2015. The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. J. Transp. Geogr. 43, 151–159. https://doi.org/10.1016/j.jtrangeo.2015.01.009.
Cornet, Y., Lugano, G., Georgouli, C., Milakis, D., 2022. Worthwhile travel time: a conceptual framework of the perceived value of enjoyment, productivity and fitness while travelling. Transp. Rev. 42 (5), 580–603. https://doi.org/10.1080/01441647.2021.1983067.
Correia, G.H., Looff, E., van Cranenburgh, S., Snelder, M., van Arem, B., 2019. On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: theoretical insights and results from a stated preference survey. Transp. Res. A Policy Pract. 119, 359–382. https://doi.org/10.1016/j.tra.2018.11.016.
Correia, G., Viegas, J.M., 2011. Carpooling and carpool clubs: clarifying concepts and assessing value enhancement possibilities through a Stated Preference web survey in Lisbon, Portugal. Transp. Res. A Policy Pract. 45 (2), 81–90. https://doi.org/10.1016/j.tra.2010.11.001.
de Dios Ortúzar, J., Willumsen, L.G., 2011. Modelling transport, 4th ed. John Wiley & Sons, Chichester, UK.
De Winter, J., van Leeuwen, P. M., Happee, R., 2012. Advantages and disadvantages of driving simulators: A discussion. In Proceedings of measuring behavior (Vol. 2012, pp. 28-31). Utrecht: sn.
De Winter, J.C., Happee, R., Martens, M.H., Stanton, N.A., 2014. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transport. Res. F: Traffic Psychol. Behav. 27, 196–217. https://doi.org/10.1016/j.trf.2014.06.016.
Detjen, H., Pfleging, B., & Schneegass, S., 2020. A Wizard of Oz field study to understand non-driving-related activities, trust, and acceptance of automated vehicles. In 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 19-29). https://doi.org/10.1145/3409120.3410662.
Dillmann, J., Den Hartigh, R.J.R., Kurpiers, C.M., Raisch, F.K., Kadrileev, N., Cox, R.F.A., De Waard, D., 2023. Repeated conditionally automated driving on the road: how do drivers leave the loop over time? Accid. Anal. Prev. 181, 106927. https://doi.org/10.1016/j.aap.2022.106927.
Ejichukwu, E., Tong, L., Hazime, G., Jia, B., 2024. Enhancing autonomous vehicle design and testing: a comprehensive review of AR and VR integration. arXiv preprint arXiv:2404.19021. Accessed https://arxiv.org/pdf/2404.19021.
Ekman, F., Johansson, M., Bligård, L.O., Karlsson, M., Strömberg, H., 2019. Exploring automated vehicle driving styles as a source of trust information. Transport. Res. F: Traffic Psychol. Behav. 65, 268–279. https://doi.org/10.1016/j.trf.2019.07.026.
Elbanhawi, M., Simic, M., Jazar, R., 2015. In the passenger seat: investigating ride comfort measures in autonomous cars. IEEE Intell. Transp. Syst. Mag. 7 (3), 4–17. https://doi.org/10.1109/MITS.2015.2405571.
Elo, S., Kyngäs, H., 2008. The qualitative content analysis process. J. Adv. Nurs. 62 (1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x.
Eriksson, A., Banks, V.A., Stanton, N.A., 2017. Transition to manual: Comparing simulator with on-road control transitions. Accid. Anal. Prev. 102, 227–234. https://doi.org/10.1016/j.aap.2017.03.011.
Eriksson, P.E., Swenberg, T., Zhao, X., Eriksson, Y., 2018. How gaze time on screen impacts the efficacy of visual instructions. Heliyon 4 (6). https://doi.org/10.1016/j.heliyon.2018.e00660.
Etzioni, S., Daziano, R.A., Ben-Elia, E., Shiftan, Y., 2021. Preferences for shared automated vehicles: a hybrid latent class modeling approach. Transp. Res. Part C Emerging Technol. 125, 103013. https://doi.org/10.1016/j.trc.2021.103013.
European Commission, 2015. Driver distraction. European Road Safety Observatory. Directorate-General for Mobility and Transport. https://road-safety.transport.ec.europa.eu/document/download/ef6ce711-f91e-49a6-b4ae-de2e133f2000_en?filename=ersosynthesis2015-driverdistraction25_en.pdf.
European Commission, 2022Driver distraction: Road safety thematic report. European Road Safety Observatory. Directorate-General for Transport. https://road-safety.transport.ec.europa.eu/document/download/662f244a-b3ad-442a-aa60-50f8437c86f2_en?filename=Road_Safety_Thematic_Report_Driver_Distraction_2022.pdf.
Fan, Q., van Essen, J.T., Correia, G.H., 2023. Optimising fleet sizing and management of shared automated vehicle (SAV) services: a mixed-integer programming approach integrating endogenous demand, congestion effects, and accept/reject mechanism impacts. Transp. Res. Part C Emerging Technol. 157, 104398. https://doi.org/10.1016/j.trc.2023.104398.
Farooq, B., Cherchi, E., Sobhani, A., 2018. Virtual immersive reality for stated preference travel behavior experiments: a case study of autonomous vehicles on urban roads. Transp. Res. Rec. 2672 (50), 35–45. https://doi.org/10.1177/0361198118776810.
Feys, M., Rombaut, E., Vanhaverbeke, L., 2021. Does a test ride influence attitude towards autonomous vehicles? a field experiment with pretest and posttest measurement. Sustainability 13 (10), 5387. https://doi.org/10.3390/su13105387.
Gerber, M.A., Schroeter, R., Xiaomeng, L., Elhenawy, M., 2020. Self-interruptions of non-driving related tasks in automated vehicles: Mobile vs head-up display. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–9. https://doi.org/10.1145/3313831.3376751.

Greifenstein, M., 2024. Factors influencing the user behaviour of shared autonomous vehicles (SAVs): a systematic literature review. Transport. Res. F: Traffic Psychol. Behav. 100, 323–345. https://doi.org/10.1016/j.trf.2023.10.027.

González, R.M., Marrero, Á.S., Marrero, G.A., 2016. How the values of travel time change when a panel data around a new tram implementation is used. Eur. J. Transp. Infrastruct. Res. 16 (4). https://doi.org/10.18757/ejtir.2016.16.4.3158.

Haghani, M., Bliemer, M.C., Rose, J.M., Oppewal, H., Lancsar, E., 2021. Hypothetical bias in stated choice experiments: Part I. Macro-scale analysis of literature and integrative synthesis of empirical evidence from applied economics, experimental psychology and neuroimaging. J. Choice Model. 41, 100309. https://doi.org/10.1016/j.jocm.2021.100309.

Hair Jr, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Danks, N.P., Ray, S., 2021. Partial least squares structural equation modeling (PLS-SEM) using R: A workbook. Springer Nature, p. 197.

Hamadneh, J., Esztergár-Kiss, D., 2022. The preference of onboard activities in a new age of automated driving. Eur. Transp. Res. Rev. 14 (1), 15. https://doi.org/10.1186/s12544-022-00540-7.

Hartwich, F., Witzlack, C., Beggiato, M., Krems, J.F., 2019. The first impression counts–a combined driving simulator and test track study on the development of trust and acceptance of highly automated driving. Transport. Res. F: Traffic Psychol. Behav. 65, 522–535. https://doi.org/10.1016/j.trf.2018.05.012.

Hernandez, J., Liu, Z., Hulten, G., DeBarr, D., Krum, K., Zhang, Z., 2013, April. Measuring the engagement level of TV viewers. In 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG) (pp. 1-7). IEEE.

Hess, S., Train, K., 2017. Correlation and scale in mixed logit models. J. Choice Model. 23, 1–8. https://doi.org/10.1016/j.jocm.2017.03.001.

IBM Corp, 2017. IBM SPSS Statistics for Windows, IBM Corp, NY, USA.

Jensen, A.F., Cherchi, E., Mabit, S.L., 2013. On the stability of preferences and attitudes before and after experiencing an electric vehicle. Transp. Res. Part D: Transp. Environ. 25, 24–32. https://doi.org/10.1016/j.trd.2013.07.006.

Johansson, M.V., Heldt, T., Johansson, P., 2006. The effects of attitudes and personality traits on mode choice. Transp. Res. Part A: Policy Practice 40 (6), 507–525. https://doi.org/10.1016/j.tra.2005.09.001.

Keseru, I., Macharis, C., 2018. Travel-based multitasking: review of the empirical evidence. Transp. Rev. 38 (2), 162–183. https://doi.org/10.1080/01441647.2017.1317048.

Keseru, I., Heyndels, E., Ton, T.D., Macharis, C., 2020. Multitasking on the go: an observation study on local public transport in Brussels. Travel Behav. Soc. 18, 106–116. https://doi.org/10.1016/j.tbs.2019.10.003.

Kettle, L., Lee, Y.C., 2022. Augmented reality for vehicle-driver communication: a systematic review. Safety 8 (4), 84. https://doi.org/10.3390/safety8040084.

Kircher, K., Ahlstrom, C., Kircher, A., 2009, June. Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test. In Driving Assessment Conference (Vol. 5, No. 2009). University of Iowa.

Klingegård, M., Andersson, J., Habibovic, A., Nilsson, E., Rydström, A., 2020. Drivers' ability to engage in a non-driving related task while in automated driving mode in real traffic. IEEE Access 8, 221654–221668. https://doi.org/10.1109/ACCESS.2020.3043428.

Ko, S.M., Ji, Y.G., 2018. How we can measure the non-driving-task engagement in automated driving: comparing flow experience and workload. Appl. Ergon. 67, 237–245. https://doi.org/10.1016/j.apergo.2017.10.009.

Kouwenhoven, M., de Jong, G., 2018. Value of travel time as a function of comfort. J. Choice Model. 28, 97–107. https://doi.org/10.1016/j.jocm.2018.04.002.

Kouwenhoven, M., Muller, J., Thoen, S., Willigers, J., de Jong, G., 2023. Values of time, reliability and comfort in the Netherlands 2022. New values for passenger travel and freight transport. (Technical report. Project number 19022).

Kyriakidis, M., Happee, R., de Winter, J.C., 2015. Public opinion on automated driving: results of an international questionnaire among 5000 respondents. Transport. Res. F: Traffic Psychol. Behav. 32, 127–140. https://doi.org/10.1016/j.trf.2015.04.014.

Lee, J., Lee, E., Yun, J., Chung, J.H., Kim, J., 2021. Latent heterogeneity in autonomous driving preferences and in-vehicle activities by travel distance. J. Transp. Geogr. 94, 103089. https://doi.org/10.1016/j.jtrangeo.2021.103089.

Levordashka, A., Stanton Fraser, D., Gilchrist, I.D., 2023. Measuring real-time cognitive engagement in remote audiences. Sci. Rep. 13 (1), 10516. https://doi.org/10.1038/s41598-023-37209-7.

Li, X., Schroeter, R., Rakotonirainy, A., Kuo, J., Lenné, M.G., 2020. Effects of different non-driving-related-task display modes on drivers' eye-movement patterns during take-over in an automated vehicle. Transport. Res. F: Traffic Psychol. Behav. 70, 135–148. https://doi.org/10.1016/j.trf.2020.03.001.

Liu, P., Jiang, Z., Li, T., Wang, G., Wang, R., Xu, Z., 2021. User experience and usability when the automated driving system fails: Findings from a field experiment. Accid. Anal. Prev. 161, 106383. https://doi.org/10.1016/j.aap.2021.106383.

Liu, P., Xu, Z., Zhao, X., 2019. Road tests of self-driving vehicles: affective and cognitive pathways in acceptance formation. Transp. Res. A Policy Pract. 124, 354–369. https://doi.org/10.1016/j.tra.2019.04.004.

Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated choice methods: analysis and applications. Cambridge University Press.

Lukovics, M., Prónay, S., Majó-Petri, Z., Kovács, P., Ujházi, T., Volosin, M., Palatinus, Z., Keszey, T., 2023. Combining survey-based and neuroscience measurements in customer acceptance of self-driving technology. Transport. Res. F: Traffic Psychol. Behav. 95, 46–58. https://doi.org/10.1016/j.trf.2023.03.016.

Malokin, A., Circella, G., Mokhtarian, P.L., 2021. Do millennials value travel time differently because of productive multitasking? a revealed-preference study of Northern California commuters. Transportation 48, 2787–2823. https://doi.org/10.1007/s11116-020-10148-2.

Meyer, J., Becker, H., Bösch, P.M., Axhausen, K.W., 2017. Autonomous vehicles: the next jump in accessibilities? Research in Transport Economics 62, 80–91. https://doi.org/10.1016/j.retrec.2017.03.005.

Milakis, D., Van Arem, B., Van Wee, B., 2017. Policy and society related implications of automated driving: a review of literature and directions for future research. J. Intell. Transp. Syst. 21 (4), 324–348. https://doi.org/10.1080/15472450.2017.1291351.

Minhas, S., Hernández-Sabaté, A., Ehsan, S., McDonald-Maier, K.D., 2020. Effects of non-driving related tasks during self-driving mode. IEEE Trans. Intell. Transp. Syst. 23 (2), 1391–1399. https://doi.org/10.1109/TITS.2020.3025542.

Mobility Innovation Centre Delft (MICD), 2022. Autonomous Driving. Retrieved from https://micd.tudelftcampus.nl/projects/autonomous-driving-passenger-experience/.

Molin, E., Adjenughwure, K., de Bruyn, M., Cats, O., Warffemius, P., 2020. Does conducting activities while traveling reduce the value of time? evidence from a within-subjects choice experiment. Transp. Res. A Policy Pract. 132, 18–29. https://doi.org/10.1016/j.tra.2019.10.017.

National Highway Traffic Safety Administration, 2012. Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices (Report No. DOT HS 811 541). https://www.nhtsa.gov/sites/nhtsa.gov/files/811541.pdf.

Naujoks, F., Befelein, D., Wiedemann, K., Neukum, A., 2018. A review of non-driving-related tasks used in studies on automated driving. In Advances in Human Aspects of Transportation: Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17− 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8 (pp. 525-537). Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_52.

Naujoks, F., Purucker, C., Neukum, A., 2016. Secondary task engagement and vehicle automation–comparing the effects of different automation levels in an on-road experiment. Transport. Res. F: Traffic Psychol. Behav. 38, 67–82. https://doi.org/10.1016/j.trf.2016.01.011.

Naujoks, F., Purucker, C., Wiedemann, K., Marberger, C., 2019. Noncritical state transitions during conditionally automated driving on German freeways: effects of non-driving related tasks on takeover time and takeover quality. Hum. Factors 61 (4), 596–613. https://doi.org/10.1177/0018720818824002.

Nazari, F., Noruzoliaee, M., Mohammadian, A.K., 2018. Shared versus private mobility: modeling public interest in autonomous vehicles accounting for latent attitudes. Transp. Res. Part C Emerging Technol. 97, 456–477. https://doi.org/10.1016/j.trc.2018.11.005.

Nibud, 2023. Autokosten [Car expenses] Retrieved from https://www.nibud.nl/onderwerpen/uitgaven/autokosten/.

Nielsen, N.H., 2021. Head Pose Estimation [Source code]. https://github.com/niconielsen32/ComputerVision/blob/master/headPoseEstimation.py.

Noble, A.M., Miles, M., Perez, M.A., Guo, F., Klauer, S.G., 2021. Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. Accid. Anal. Prev. 151, 105959. https://doi.org/10.1016/j.aap.2020.105959.

Nordhoff, S., De Winter, J., Kyriakidis, M., Van Arem, B., Happee, R., 2018a. Acceptance of driverless vehicles: results from a large cross-national questionnaire study. J. Adv. Transp. 2018. https://doi.org/10.1155/2018/5382192.

Nordhoff, S., de Winter, J., Madigan, R., Merat, N., van Arem, B., Happee, R., 2018b. User acceptance of automated shuttles in Berlin-Schöneberg: a questionnaire study. Transport. Res. F: Traffic Psychol. Behav. 58, 843–854. https://doi.org/10.1016/j.trf.2018.06.024.

Osz, K., Rydström, A., Fors, V., Pink, S., Brostrom, R., 2018. Building collaborative test practices: design ethnography and WOz in autonomous driving research. Interaction Design and Architecture(s) Journal – Ixd&a 37, 12–20. https://doi.org/10.55612/s-5002-037-001.

Öztürker, M., de Almeida, H., Correia, G., Scheltes, A., Olde Kalter, M.J., van Arem, B., 2022. Exploring users' preferences for automated minibuses and their service type: a stated choice experiment in the Netherlands. J. Adv. Transp. 2022, 4614848. https://doi.org/10.1155/2022/4614848.

Pakusch, C., Stevens, G., Boden, A., Bossauer, P., 2018. Unintended effects of autonomous driving: a study on mobility preferences in the future. Sustainability 10 (7), 2404. https://doi.org/10.3390/su10072404.

Pech, T., Enhuber, S., Wandtner, B., Schmidt, G., Wanielik, G., 2019. Real time recognition of non-driving related tasks in the context of highly automated driving. In Advanced Microsystems for Automotive Applications 2018: Smart Systems for Clean, Safe and Shared Road Vehicles 22nd (pp. 43-55). Springer International Publishing. https://doi.org/10.1007/978-3-319-99762-9_4.

Pudāne, B., Molin, E.J., Arentze, T.A., Maknoon, Y., Chorus, C.G., 2018. A time-use model for the automated vehicle-era. Transp. Res. Part C: Emerging Technol. 93, 102–114. https://doi.org/10.1016/j.trc.2018.05.022.

Pudāne, B., Rataj, M., Molin, E.J., Mouter, N., van Cranenburgh, S., Chorus, C.G., 2019. How will automated vehicles shape users' daily activities? Insights from focus groups with commuters in the Netherlands. Transp. Res. Part D: Transp. Environ. 71, 222–235. https://doi.org/10.1016/j.trd.2018.11.014.

Pudāne, B., Van Cranenburgh, S., Chorus, C.G., 2021. A day in the life with an automated vehicle: empirical analysis of data from an interactive stated activity-travel survey. J. Choice Model. 39, 100286. https://doi.org/10.1016/j.jocm.2021.100286.

Riegler, A., Riener, A., Holzmann, C., 2021. A systematic review of virtual reality applications for automated driving: 2009–2020. Front. Hum. Dyn. 3, 689856. https://doi.org/10.3389/fhumd.2021.689856.

Rizki, M., Joewono, T.B., Dharmowijoyo, D.B., Belgiawan, P.F., 2021. Does multitasking improve the travel experience of public transport users? investigating the activities during commuter travels in the Bandung Metropolitan Area, Indonesia. Public Transp. 13, 429–454. https://doi.org/10.1007/s12469-021-00263-3.

Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48, 1–36. http://www.jstatsoft.org/v48/i02/.

SAE International, 2021. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Retrieved from.

Seaman, S., Gershon, P., Angell, L., Mehler, B., Reimer, B., 2022. Non-driving-related task engagement: the role of speed. Safety 8 (2), 34. https://doi.org/10.3390/safety8020034.

Shahini, F., Zahabi, M., 2022. Effects of levels of automation and non-driving related tasks on driver performance and workload: a review of literature and meta-analysis. Appl. Ergon. 104, 103824. https://doi.org/10.1016/j.apergo.2022.103824.

Shi, E., Bengler, K., 2022. Non-driving related tasks' effects on takeover and manual driving behavior in a real driving setting: a differentiation approach based on task switching and modality shifting. Accid. Anal. Prev. 178, 106844. https://doi.org/10.1016/j.aap.2022.106844.

Singleton, P.A., 2019. Discussing the "positive utilities" of autonomous vehicles: will travellers really use their time productively? Transp. Rev. 39 (1), 50–65. https://doi.org/10.1080/01441647.2018.1470584.

Singleton, P.A., 2020. Multimodal travel-based multitasking during the commute: who does what? Int. J. Sustain. Transp. 14 (2), 150–162. https://doi.org/10.1080/15568318.2018.1536237.

Skippon, S.M., Kinnear, N., Lloyd, L., Stannard, J., 2016. How experience of use influences mass-market drivers' willingness to consider a battery electric vehicle: a randomised controlled trial. Transp. Res. A Policy Pract. 92, 26–42. https://doi.org/10.1016/j.tra.2016.06.034.

Solís-Marcos, I., Ahlström, C., Kircher, K., 2018. Performance of an additional task during Level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation. Hum. Factors 60 (6), 778–792. https://doi.org/10.1177/0018720818773636.

Sun, S., Wong, Y.D., 2022. Travel-based multitasking behaviour in Singapore: determinants and impacts on money-time-seat trade-offs. Travel Behav. Soc. 26, 84–95. https://doi.org/10.1016/j.tbs.2021.09.005.

Susilo, Y.O., Lyons, G., Jain, J., Atkins, S., 2012. Rail passengers' time use and utility assessment: 2010 findings from Great Britain with multivariate analysis. Transp. Res. Rec. 2323 (1), 99–109. https://doi.org/10.3141/2323-12.

Train, K.E., 2009. Discrete choice methods with simulation. https://eml.berkeley.edu/books/choice2.html.

van Gent, P., Farah, H., Nes, N. V., & van Arem, B., 2017. Towards Real-Time, Nonintrusive Estimation of Driver Workload in a Simulated Environment. In Proceedings of the 6th International Conference on Road Safety & Simulation (RSS): 17-19 October 2017, The Hague, Netherlands. Article 283. https://research.tudelft.nl/files/31672191/Towards_Real_Time_Nonintrusive2.pdf.

Varghese, V., Jana, A., 2018. Impact of ICT on multitasking during travel and the value of travel time savings: empirical evidences from Mumbai, India. Travel Behav. Soc. 12, 11–22. https://doi.org/10.1016/j.tbs.2018.03.003.

Wadud, Z., Huda, F.Y., 2019, February. Fully automated vehicles: the use of travel time and its association with intention to use. In Proceedings of the Institution of Civil Engineers-Transport (Vol. 176, No. 3, pp. 127-141). Thomas Telford Ltd.

Wardman, M., 2004. Public transport values of time. Transp. Policy 11 (4), 363–377. https://doi.org/10.1016/j.tranpol.2004.05.001.

Wardman, M., Chintakayala, P., Heywood, C., 2020. The valuation and demand impacts of the worthwhile use of travel time with specific reference to the digital revolution and endogeneity. Transportation 47 (3), 1515–1540. https://doi.org/10.1007/s11116-019-10059-x.

Weidner, F., Hoesch, A., Poeschl, S., Broll, W., 2017. Comparing VR and non-VR driving simulations: an experimental user study. In: 2017 IEEE Virtual Reality (VR), pp. 281–282. https://doi.org/10.1109/VR.2017.7892286.

Yang, L., Wang, R., Zhao, X., Xu, Z., Yang, Y., 2021. CAVTest: a closed connected and automated vehicles test field of Chang'an University in China. SAE Int. J. Connected Automated Vehicles 4 (12–04-04-0032), 423–435. https://doi.org/10.4271/12-04-04-0032.

Yap, M.D., Correia, G., Van Arem, B., 2016. Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. Transp. Res. A Policy Pract. 94, 1–16. https://doi.org/10.1016/j.tra.2016.09.003.

Yu, H., Gupta, A., Lee, W., Arroyo, I., Betke, M., Allesio, D., et al., 2021, July. Measuring and integrating facial expressions and head pose as indicators of engagement and affect in tutoring systems. In International Conference on Human-Computer Interaction (pp. 219-233). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-77873-6_16.

Zou, X., O'Hern, S., Ens, B., Coxon, S., Mater, P., Chow, R., Neylan, M., Vu, H.L., 2021. On-road virtual reality autonomous vehicle (VRAV) simulator: an empirical study on user experience. Transp. Res. Part C Emerging Technol. 126, 103090. https://doi.org/10.1016/j.trc.2021.103090.