2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 979-8-3503-4839-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/VRW58643.2023.00023

Towards a 3D Evaluation Dataset for User Acceptance of Automated Shuttles

Ming Yan*
Politecnico di Milano

Wei Geng†
The Hong Kong University of Science and Technology (Guangzhou)

Pan Hui‡
The Hong Kong University of Science and Technology (Guangzhou)

Abstract

The popularity of automatic driving technology has gradually freed users from driving tasks and increased interaction with vehicles and machines. Understanding user acceptance and making them more receptive to new technologies can help businesses and researchers find better ways to design Human-Machine Interactions. The simulation experiment in an immersion environment can evaluate the user's acceptance of the design with low cost and high efficiency. Further, the evaluation methods of some existing studies are different, which creates obstacles to the reuse and reference of research results between different scholars. However, there are limited simulation data that can be used for such interactive evaluation, such as typical 3D environment data based on Virtual Reality devices. We design dataset, an ongoing 3D test dataset produced by Unity software, to be employed by different studies to evaluate interaction design for autonomous driving. The physical medium, composition, test participants, and procedure of the 3D environment data are described in this paper.

Index Terms: Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies; Human-centered computing—Interaction design—Empirical studies in interaction design

1 Introduction

The maturation of autonomous driving (AD) technology heralds the future of large-scale commercialization, the currently existing underground AD serves as an example [6, 10]. Shared Autonomous Vehicles (SAV), especially automated shuttles, are in the spotlight. This type of vehicle of level 4 in the range of 0 to 5 [2] can handle most traffic situations without a human attendant. Modern vehicles already integrate numerous digital interactive devices such as touch panels, microphones, cameras, GPS sensors, and light sensors. These facilities will broaden the variety of possible Human-Machine Interactions (HMI), which have gradually changed from driving as the main task to leisure and entertainment as the primary purpose. The user experience of HMI in Autonomous Vehicles (AVs) will become more critical in such a scenario.

It is reported that the biggest obstacle to the large-scale adoption of AVs would be psychological factors, such as acceptance and trust, rather than technical issues [1, 19]. Direct physical design to address the issues above would be costly and time-consuming. In order to guide the design and evaluation of HMI in AVs, it is crucial to create a simulation platform based on Virtual Reality (VR) for people with no experience with AVs. An immersive 3D environment can be employed to assess user experience better and get insights from it. However, according to our best knowledge, there exists no standardized 3D simulation test environment for the evaluation of HMI in AVs. Therefore, this work mainly introduces the 3D simulation environment data we are developing, which will be published as a standard testset and can be adopted by other peers.

We first analyzed all aspects of user acceptance of AV and designed a targeted evaluation experiment process. This testing process is mainly carried out in the created 3d environment using VR. Our 3D data will mainly consist of different elements affecting user experience and acceptance namely the appearance of AV, external interactive signal equipment, interactive equipment inside the AV, scenarios, weather, scenery, fixed routes, stations, and voice packages. We further specify the testing procedure on this dataset and what result will be collected to assess user experience and acceptance of AVs.

At present, the evaluation methods for the interactive experience and user acceptance of AD are not clear and unified. A standard assessment environment and method can promote the research progress in this area. Our ongoing work aims to create a standard 3D simulation environment testset and define a consistent testing procedure using this data. This standardized testset can be easily adopted by different design developers, and its result can be utilized for comparison and re-design. The remainder of this paper is organized as follows:

- Section 2 presents the research background, related concepts, and findings.
- Section 3 describes the 3D dataset, including the test tool, requirements, and evaluation processes based on the dataset.
- Section 4 discusses the limitations and envisions future work.
- Section 5 summarizes and concludes the article.

2 Background

Automated shuttle is a specially constructed vehicle for safely transferring human passengers or goods from point A to point B and a shared service across multiple communities of target customers [5, 11]. Although a certain number of studies apply user acceptance models to statistically analyze users' experience data, the majority of research on this topic evaluates user acceptability by evaluating respondents' usage and purchase intentions, with an emphasis on subjective data collecting. However, subjective evaluations vary significantly in different situations, which will bring some bias in the results. This project's primary objective is to design as standard a test dataset and process as possible to obtain more objective results.

2.1 Method

This section is divided into the following four aspects:

1. Summarize the current user acceptance evaluation analysis methods and update the evaluation model based on the original ones.

*e-mail: ming.yan@polimi.it
†e-mail: wgeng252@connect.hkust-gz.edu.cn
‡e-mail: panhui@ust.hk

979-8-3503-4839-2/23/$31.00 ©2023 IEEE
DOI 10.1109/VRW58643.2023.00023

2. Propose the design principle for the interior and exterior HMI test dataset development for automated shuttles based on a review of the academic and industrial efforts.

3. Communicate the design concept through prototyping and simulation. Specifically, creating immersive VR simulation datasets to describe the interactions within a trip situation.

4. Introduce the particular testing procedure that will be used to evaluate the test dataset in subsequent work and is explained in section 2.2.

2.1.1 User Acceptance Evaluation Model

"User acceptance" is a perspective of user research, which is used to assess potential consumers' overall attitudes and behavioral responses toward new technologies [18]. This study defines "acceptance" as potential users' positive attitude regarding AV technology before trying it [14]. Academics have proposed various models for evaluating user acceptance. Among them, the Unified Theory of Acceptance and Use of Technology (UTAUT) model [8] and the Car Technology Acceptance Model (CTAM) proposed by Osswald et al. [17, 23] are the most popular. Because neither model is specialized to the HMI elements of autonomous driving, there is no validated testset to assist in the development of user acceptability models in this domain.

By classifying the relevant models, we were able to construct a new model that updates the Model of Automated Vehicle Acceptance (MAVA) [13] with innovations and adjustments. It defines external variables, a process model with main impacts, and individual differences (such as socio-demographics, travel behavior, and personality) at the macro, meso, and micro levels, with the dataset for this study concentrating on the four-stage decision process at the meso level. First and foremost, access to AVs. Second, the development of favorable attitudes toward self-driving vehicles. Third is the decision to use an autonomous vehicle, and finally, the actual use of AVs.

2.1.2 Design Principles and Prototyping for Test Dataset

In order to design a 3D dataset of HMI systems in automated shuttles for testing, it is necessary to investigate technology trends in automotive HMI design via a comprehensive review and a structured analysis of the academic and commercial perspectives, including concept shuttles. To examine the utilization of interactive technology and infotainment features envisioned for future travel scenarios, it is required to undertake a thorough examination of selected shuttles.

Passengers are not obliged to operate a vehicle while using public transportation [20]. The HMI design for automated shuttles focuses on supporting non-driving-related activities (NDRTs) inside the vehicle and information exchange of HMI outside the vehicle. The inside HMI system gradually transitions from a single central screen to multiple screens in various positions and sizes as the layout and seating arrangements alter. Its carrier also consists of physical buttons display panels, and applications that may be used to get information on shuttle bus operations, scheduled routes, etc. The External HMI (eHMI) equips the front and back of the shuttles with displays for text, expressions, and other indications. However, it is worth investigating how to avoid information overload while grabbing users' attention through suitable engagement. Consequently, our testing procedure recommends building two experimental datasets in Unity 3D for interior and exterior scenarios by requiring users to accomplish various interactive operating tasks. For instance, Unity 3D can be used to create a virtual display screen within the shuttle to aid participants in responding to information requests, or different interface devices can be created outside the vehicle to assess the user's communication efficiency.

2.2 Introduction to Testing Protocol

Test Protocol is a collection of Test Cases that check a specific system element. Each test case consists of multiple test stages. The subject must be tested in a standardized process during the initial design phase and later iterations. This paper defines a methodological procedure to observe users' behavior and subjective evaluations during specified cases and conditions, which is organized into three macro phases, presented in chronological order: (1) Preliminary preparation; (2) test execution; (3) data aggregation. The preparatory preparation, as outlined in Sections 3.1 and 3.2, entails confirming the equipment, environment, regulations, and choices of the tested personnel, etc. In addition, test execution denotes the procedure and instruction in the test, which is listed in section 3.3. for data aggregation.

3 3D Test Dataset

In other areas where VR has made some progress as a simulation testing tool, we describe the advantages of VR over existing evaluation methods. Our 3D test dataset is subsequently defined in detail. Finally, based on this test dataset, some test processes that can be carried out are introduced.

3.1 VR as a Tool for Simulation

It is challenging to assess user acceptance of AVs used in public transport because of the lack of prototypes [22]. Implementing such projects requires systematic support from the government, transportation system, urban road planning departments, etc. [4]. The user research should be carried out before the official launch of the automated shuttles project because of the manufacturing cost. Therefore, evaluating them requires the use of techniques other than physical prototypes.

The methods commonly used in existing research are divided into structured questionnaires [3], Wizard-of-Oz experiments [15], and VR experiments [**?**]. The limitation of structured questionnaires is that it is based on participant responses rather than their actual behaviors. Due to the user's lack of imagination and experience, their responses may differ from actual conduct. As for Wizard-of-Oz experiments, although the illusion of AD scenarios can be simulated under the operation of the staff, the scenes still need to be carried out in actual road conditions, which leads to technical safety risks. In addition, existing related studies are mainly aimed at private cars. It is not easy for this method to realize the scenario simulation in a public transportation because of the diversity and changeability in the transit system's environment.

On the contrary, VR technology has strong applicability in investigating users' acceptance of emerging technologies. Rebelo et al. summarize the advantages of using VR for user experience research into three topics: availability, safety, and data provision [16]. Availability here refers to simulating specific contexts in a repeatable and systematic manner without spending the time and cost required by the real/physical setups. Safety is not only about preventing injuries but also enables practices through trials and errors without being literally affected by their social-practical consequences. What is meant by data provision is to help researchers collect data even in the initial stages of the design process confirming "high accuracy and good ecological validity". By contrast, VR also has its disadvantage: the risk of simulation sickness symptoms. Using Oculus Rift Head-Mounted Displays (HMDs) to examine interactions with motorized or non-motorized interaction partners (cars, motorcycles, cyclists) and human users inside and outside the vehicle would immerse users in the virtual environment easier.

As a result, for the test environment and apparatus, we recommend coupling with a Leap Motion controller as an input device that digitizes the hands of the end-user in real-time. The Leap Motion Application Programming Interface (API) includes an effective skeletal tracking model that provides additional information about hands and fingers and improves overall tracking data. In this way, users can select the part of their interest on the virtual object on the screen. The software application for managing the VR devices should be implemented using Unity 3D, a game engine widely used to develop games and interactive VR environments.

3.2 Design of the Test Dataset

3.2.1 Test Environment and Apparatus

The location for the experimental protocol should be a controlled environment while using an automated shuttle as the prototype, which should drive in the virtual environment according to specified routes. The driving simulator used for the test should include a realistic driver seat, which provides a surrounding audio system. It is more economical than the actual scenario, which is suitable for the early design shown in figure 1.

Figure 1: Driving Simulator Setup for the test dataset [7].

Test Dataset for Interior HMI Most recent research on this direction in the HMI field is devoted to several perspectives, namely 1) Transparency of communication, 2) Satisfaction of classified user groups' needs, 3Interior development for different scenes, 4Remote Operation for real-time management, and monitoring, 5Control transfer in an emergency. No matter which aspects, the in-vehicle HMIs should support the passengers' tasks of supervising the driving environment when needed and self-regulating their non-driving related activities (NDRAs). Such support may be provided by either continuously presenting information on automation reliability. The scenario built in Unity 3D was a riding experience inside an autonomous shuttle bus driving around a specified area (Fig. 2).

Figure 2: A representation of the Unity 3D interaction scene was created for testing inside the vehicle. There are two interactive displays in the scene, one above the user's head and the other on the armrest rail.

Test Dataset for Exterior HMI As for the application scenarios of HMI outside the bus, it should be considered whether it can support the interactive objects outside to complete faster decision-making for crossing the street and prevent their misperceptions and behaviors to a greater extent. Interactive objects here refer to pedestrians, motor vehicles, non-motor vehicles, other AVs, Etc. When there is no conflict between the bus's driving direction and the interactive objects outside the car that need attention, whether the bus is autonomous has no substantial impact on other interactive objects, so these scenarios will not affect the user's acceptance of the technology. In 2020, as summarized by Christina et al., there are three scenarios in which the two objects have conflicts [9], as figure 3 illustrates namely 1) The interaction partner approaches the automated vehicle frontally, 2) Orthogonally from the side, 3) Merges in front of the automated vehicle with a lateral approach direction.

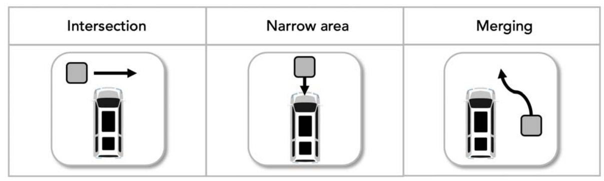

Figure 3: The three conflicting situations should be included in the testing environment [9].

The test dataset for eHMI will be executed for the above three scenarios. In addition, the test procedure should be carried out in a traffic environment without right-of-way rules when eHMI is the only standard. In this scenario, the Vulnerable Road Users (VRUs) facing the eHMI cannot control the bus but receive the information transmitted, affecting its decision-making. The scenario built in Unity 3D was proposed in figure 4.

Figure 4: A representation of the Unity 3D interaction scene was created for testing outside the vehicle. The test dataset's environment display (left); external interactive signal equipment (right).

3.2.2 Participants

The sample size should be large enough to draw reliable conclusions from experimental data and involve at least 20 subjects. The target population interacting with automated shuttles in the future is vast. Therefore, there are no explicit restrictions on people (all ages, nationalities, education levels, heights, etc.) eligible to take the test. To obtain a representative age distribution, follow the proposal proposed by the National Highway Traffic Safety Administration (NHTSA) in the USA, different age groups of n = 5 each, 18–24, 25–39, 40–54, and older than 54 [12].

Participants need to meet further prerequisites. While this experiment aims to explore test environments in levels 4 and 5 that support the unmanned operation of vehicles, participants are still required to be familiar with, or experience assisted driving technologies (i.e., lane-keeping systems and adaptive cruise control).

Besides, preferably have a valid driver's license. At the same time, the participation criteria were not to be over-sensitive to activities that might create motion sickness. The participants preferably had previously experienced VR-HMD, but no experience in this area still be allowed. All participants had a 5-minute warm-up session before the demo to get used to the interactions.

3.3 Evaluation based on the Dataset

The evaluation procedures and instructions have been defined, including the testing process and the data to be collected.

3.3.1 Testing Process of Interior and Xxterior HMI

The experiment simulation was combined with observations, questionnaires, and participant interviews. After a briefing and consent agreement, the participants were equipped with a VR-HMD. The VR setup consisted of an empty tracked area of 4.0 m × 4.0 m with a chair to sit down on (when the participants decided to sit during the test). Besides the HMD, the participants were equipped with noise-canceling headphones. Specifically, a tutorial first familiarized the participants with the VR technology. Afterward, the test scenario was conducted in a randomized order for each HMI concept. Video-recorded observations gathered the participants' behaviors and reactions (e.g., being surprised or amused) to the events. At the end of the experience simulation in VR, the participants were requested to fill out a questionnaire regarding their perception of the HMI concepts using a five-point Likert scale [11]. Finally, the researchers conducted interviews to get insights into subjective justifications from the participants.

For the interior HMI, the simulation was created in the following eight specific tasks in the VR environment: (1) Start of scenario; (2) Confirm the destination; (3) Homepage exploration; (4) Visualize the itinerary information; (5) Visualize the vehicle's driving state; (6) Visualize the entertainment information; (7) Vehicle control transfer; (8) End of scenario.

For the exterior HMI, the perspective of the tester's experience is the pedestrians outside the bus, which can be divided into two processes. On the one side, the experiment of the first process aims to compare the difference in user experience and acceptance between ordinary vehicles driven by human drivers and automated shuttles with eHMI. The specific schematic diagram of the scenarios and the tasks that should be included in the test shown in Fig. 5 includes various scenes from A to E. On the other side, the main goal of the second process was to compare the selected HMI concept via different visualization technologies (i.e., laser projection on street/sidewalk/zebra crossing or display-based) with the control group by asking the participants to navigate through the intersections between A and B constantly.

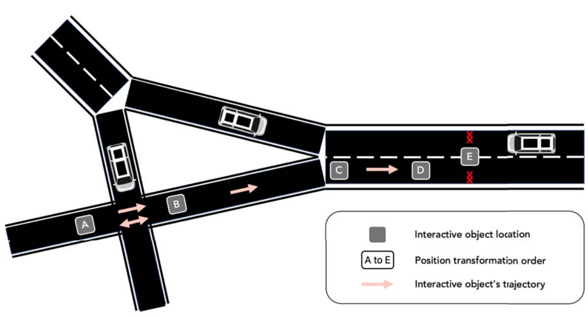

Figure 5: A diagram depicting the specific scenarios and tasks of the entire test. The testing procedure is followed from point A to point E.

3.3.2 Data Aggregate and Analysis

To provide valuable indications, quantitative and qualitative data should be collected.

Qualitative data describes users' perception and experience of HMI, which is relatively subjective feedback. Such data are usually obtained through questionnaire surveys, group interviews, on-site observation, and other methods. During the test process, the thinking-aloud practice procedure asks the participant to say what they think during the task execution, which aim is to explain their comments on existing HMI and help to understand their actions and reactions in case of unexpected interface behaviors. The usability testing will be concluded with a semi-structured interview that examines the participant's public transportation experiences and their expectations for an automated vehicle or shuttle bus.

Quantitative data can be determined by measuring and collecting users' physiological parameters and performance measures of specified tasks, which can objectively reflect users' status. For physiological data, collecting the Galvanic Skin Response (GSR), the heart rate (HR), and heart rate variability (HRV) is standard. In addition, using an eye tracker to record the position, sequence, and time of the eye-fixation interface can also help analyze the rationality of the design. Data collection related to testing tasks could be compared by recording task time, the number of clicks, and the number of successes and failures.

4 Future Work

In the future, more vehicle companies will focus on the user's emotional experience [21] because the essence of HMI is to make communication between humans and machines smoother, especially in fully automatic vehicles without the driver as a medium. AV design research will focus on users' perceptions, information classification, and HMIs. This work has defined theoretical solutions and is under development at present. The next step is to finish the 3D test dataset in Unity 3D and apply the test procedure to practice with different HMI design variants and specifications of the AD system. While doing the experiment, the 3D test dataset will also be optimized and modified according to the experimental feedback. Furthermore, the current research is mainly conducted in public transit scenarios, and the attention to specific procedures and vulnerable users is low. Future research should be more comprehensive, such as comparing the impact of different kinds of vehicles and involving different user groups in extreme weather and road conditions to evaluate the effectiveness of different HMIs from the perspective of inclusive design.

5 Conclusion

In this paper, we summarized the current research on the evaluation of user experience and acceptance of HMI and discussed their limitations. The lack of standard test scenarios makes it difficult to compare and follow related research, so the standard test process cannot be carried out. We described our work in progress which creates a 3D scene in Unity as a test dataset for evaluating user experience and acceptance of AV. The test process based on this test dataset was also envisioned. The datasets we are creating will be made publicly available, which will not only reduce duplication of workload in this field of research but also enable comparison of related assessments. In addition to completing the creation of the dataset, we also put forward several directions and priorities for future work.

Acknowledgments

The present research was conducted in the i.Drive (Interaction of Driver, Road, Infrastructure, Vehicle, and Environment) Laboratory of Politecnico di Milano (http://www.idrive.polimi.it/). This work was partially sponsored by the China Scholarship Council.

References

[1] C. G. Burns, L. Oliveira, P. Thomas, S. Iyer, and S. Birrell. Pedestrian decision-making responses to external human-machine interface designs for autonomous vehicles. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 70–75. IEEE, 2019.

[2] J. C. De Winter, R. Happee, M. H. Martens, and N. A. Stanton. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation research part F: traffic psychology and behaviour*, 27:196–217, 2014.

[3] S. M. Faas and M. Baumann. Yielding light signal evaluation for self-driving vehicle and pedestrian interaction. In *International Conference on Human Systems Engineering and Design: Future Trends and Applications*, pp. 189–194. Springer, 2019.

[4] F. Golbabaei, T. Yigitcanlar, and J. Bunker. The role of shared autonomous vehicle systems in delivering smart urban mobility: A systematic review of the literature. *International Journal of Sustainable Transportation*, 15(10):731–748, 2021.

[5] M. J. Hallewell, N. Hughes, D. R. Large, C. Harvey, J. Springthorpe, and G. Burnett. Deriving personas to inform hmi design for future autonomous taxis: A case study on user requirement elicitation. *Journal of Usability Studies*, 17(2), 2022.

[6] J. Hu, P. Bhowmick, F. Arvin, A. Lanzon, and B. Lennox. Cooperative control of heterogeneous connected vehicle platoons: An adaptive leader-following approach. *IEEE Robotics and Automation Letters*, 5(2):977–984, 2020.

[7] i.Drive Lab. Interaction between driver, road infrastructure, vehicle, and environment.

[8] A. Johnsen, N. Strand, J. Andersson, C. Patten, C. Kraetsch, and J. Takman. Literature review on the acceptance and road safety, ethical, legal, social and economic implications of automated vehicles. 2017.

[9] C. Kaß, S. Schoch, F. Naujoks, S. Hergeth, A. Keinath, and A. Neukum. Standardized test procedure for external human–machine interfaces of automated vehicles. *Information*, 11(3):173, 2020.

[10] H. S. M. Lim and A. Taeihagh. Autonomous vehicles for smart and sustainable cities: An in-depth exploration of privacy and cybersecurity implications. *Energies*, 11(5):1062, 2018.

[11] M. Liu, J. Wu, C. Zhu, and K. Hu. Factors influencing the acceptance of robo-taxi services in china: An extended technology acceptance model analysis. *Journal of Advanced Transportation*, 2022, 2022.

[12] F. Naujoks, S. Hergeth, K. Wiedemann, N. Schömig, Y. Forster, and A. Keinath. Test procedure for evaluating the human–machine interface of vehicles with automated driving systems. *Traffic injury prevention*, 20(sup1):S146–S151, 2019.

[13] S. Nordhoff, M. Kyriakidis, B. Van Arem, and R. Happee. A multi-level model on automated vehicle acceptance (mava): A review-based study. *Theoretical Issues in Ergonomics Science*, 20(6):682–710, 2019.

[14] C. Pigeon, A. Alauzet, and L. Paire-Ficout. Factors of acceptability, acceptance and usage for non-rail autonomous public transport vehicles: A systematic literature review. *Transportation research part F: traffic psychology and behaviour*, 81:251–270, 2021.

[15] C. Ranasinghe, K. Holländer, R. Currano, D. Sirkin, D. Moore, S. Schneegass, and W. Ju. Autonomous vehicle-pedestrian interaction across cultures: towards designing better external human machine interfaces (ehmis). In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2020.

[16] F. Rebelo, P. Noriega, E. Duarte, and M. Soares. Using virtual reality to assess user experience. *Human factors*, 54(6):964–982, 2012.

[17] S. Reig, S. Norman, C. G. Morales, S. Das, A. Steinfeld, and J. Forlizzi. A field study of pedestrians and autonomous vehicles. In *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*, pp. 198–209, 2018.

[18] G. Schuitema, L. Steg, and M. van Kruining. When are transport pricing policies fair and acceptable? *Social Justice Research*, 24(1):66–84, 2011.

[19] A. Shariff, J.-F. Bonnefon, and I. Rahwan. Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour*, 1(10):694–696, 2017.

[20] A. M. Tinga, D. Cleij, R. J. Jansen, S. van der Kint, and N. van Nes. Human machine interface design for continuous support of mode awareness during automated driving: An online simulation. *Transportation research part F: traffic psychology and behaviour*, 87:102–119, 2022.

[21] D. Vrščaj, S. Nyholm, and G. P. Verbong. Is tomorrow's car appealing today? ethical issues and user attitudes beyond automation. *AI & SOCIETY*, 35(4):1033–1046, 2020.

[22] M. Yan, L. Rosa Elena Rampino, H. Zhao, G. Caruso, et al. Implications of human-machine interface for inclusive shared autonomous vehicles. In *Human Factors in Transportation. AHFE (2022) International Conference. AHFE*, vol. 60, pp. 542–550. . AHFE International, 2022.

[23] J. Zmud, I. N. Sener, J. Wagner, et al. Consumer acceptance and travel behavior: impacts of automated vehicles. Technical report, Texas A&M Transportation Institute, 2016.