2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Evaluating Multimodal Communication Methods for Autonomous Buses in Pedestrian-Dense University Environments

Abdalla Ahmed Roshdi Mohamed[1*], Ashita Ashok[1], Qazi Hamza Jan[1], Franziska Babel[2], and Karsten Berns[1]

***Abstract*— In urban pedestrian zones where autonomous vehicles (AVs) increasingly operate alongside humans, clear communication between AVs and pedestrians is essential for safety and trust. This study conducted an exploratory research on pedestrian reactions to an autonomous shuttle bus (AutoBus) operating on a university campus. Using a real-world deployment, the effectiveness of visual and auditory communication cues in a real-world setting was evaluated. The AutoBus continuously looped a 480-meter path on campus during lunchtime, and pedestrians who walked toward and crossed the bus were invited to complete an online survey following this interaction. Data was collected from 58 participants at a technical university through behavioral observations and post-interaction surveys. The results reveal that visual cues were more consistently recognized than auditory ones, influencing pedestrian awareness and response. Trust in AV's safety was shaped more by its perceived safety than by prior experiences with the AutoBus. Moreover, willingness to yield was positively associated with the perceived social status of the AV, but not whether it was perceived as an autonomous robot or as representing its passengers. These findings offer practical insights for improving AV communication design to support safer, more intuitive interactions in shared spaces.**

Keywords—Field Robots, Autonomous Agents, Social Human-Robot Interaction

I. INTRODUCTION

Autonomous vehicles (AVs) are projected to have a significant road presence, with a $21 billion market by 2035 [1]. As AVs integrate into urban environments, their interactions with existing road users drive research on human-vehicle interaction (HVI) [2]. AVs serve as a catalyst for sustainable urban mobility, optimizing traffic flow, reducing emissions, and enhancing safety, particularly in structured environments with clear road signs and markings [3]. However, their deployment in pedestrian-dense spaces like campuses and shared urban areas remains challenged by public acceptance and the need for intuitive HVI [4]. In real, unstructured environments, AVs require robust communication strategies to replace traditional human driving cues like eye contact and gestures [5].

Pedestrian-AV interactions introduce unique challenges, especially in shared spaces without formal traffic infrastructure [6], [7]. Pedestrians rely heavily on implicit cues when interacting with conventional vehicles [8], but the absence of human drivers in AVs creates a significant communication gap [9]. Addressing this, Prédhumeau *et al.* [10] proposed a framework for analyzing pedestrian behavior, emphasizing that AVs must adapt to ambiguous social signals in shared spaces.

Despite prior research on AV acceptance [11], [12], few studies integrate self-reported preferences (surveys) with behavioral metrics (posture, gaze, facial expressions) in real-time pedestrian-AV interactions. Our work bridges this gap by analyzing pedestrian responses (e.g., smiles, avoidance, noticing the bus while distracted) through skeleton data [13]. This aligns with research on external Human-Machine Interfaces (eHMIs) to enhance AV legibility and trust [14], [15]. While standardized communication modalities are emphasized in prior work [16], [17], pedestrians' awareness of auditory and visual cues of autonomous buses in real-world environments is not sufficiently explored yet.

To examine these factors, our study investigates participants' reactions to an autonomous bus (AutoBus)[1] operating in a university pedestrian zone, utilizing multimodal eHMIs (auditory and visual cues) to interact with pedestrians [13]. Although the bus currently operates at walking speed as part of an experimental testing phase, the long-term objective is to support autonomous mobility for populations such as senior citizens rather than to optimize immediate travel efficiency. A university campus, therefore, provides an ideal test environment due to:

1. Mixed-use paths with frequent AV-pedestrian interactions.
2. Diverse users (students, faculty, visitors) of varying ages.
3. Varied familiarity, allowing analysis of first-time vs. repeated exposure to AVs.

II. RELATED WORK

A. Communication Modalities of Autonomous Vehicles

AVs rely on a combination of implicit cues (vehicle motion, speed) and explicit signals from eHMIs [18], [17] to communicate their intent. Research suggests that using both enhances pedestrian understanding and trust, reducing uncertainty in shared spaces [19], [20]. Dey *et al.* [15] further demonstrated that well-designed eHMIs, specifically those that clearly communicate the vehicle's intent (e.g., yielding) through both explicit visual text and implicit vehicle behavior (e.g., deceleration), significantly improve pedestrian trust and acceptance of AVs, underscoring the importance of clear, multimodal communication strategies. This was also shown for a reduction of crossing times in front of an autonomous vehicle [20].

---

[1] Department of Computer Science, RPTU Kaiserslautern-Landau, Kaiserslautern, Germany

[2] Department of Computer & Information Science, Linköping University, Sweden

`*correspondence: a_mohame@rptu.de`

[1] https://rrlab.cs.rptu.de/en/robots/autobus

979-8-3315-8160-2/26/$31.00 ©2026 IEEE

Fig. 1. Autonomous Shuttles within Campus Environment: (left to right) SMART, EZ10 Gen2, AutoBus, & Mcity.

B. Visual and Auditory Interfaces

Visual and auditory communication systems play complementary roles in enhancing AV-pedestrian interaction. Visual cues, such as LED panels and light strips, effectively convey vehicle intentions like "yielding" or "proceed" [21], while symbol-based communication reduces language barriers and increases accessibility [22]. However, visual cues often require prior knowledge for correct interpretation [23].

Auditory communication addresses these limitations by providing real-time alerts [24]. Well-timed auditory cues and adaptive warning sounds enhance pedestrian response in high-risk situations [25], [26]. Additionally, cultural differences influence auditory interface design, necessitating customizable sound profiles to ensure effective communication across diverse user groups [27]. Combining these modalities optimizes AV legibility, reduces ambiguity, and improves pedestrian safety in shared environments.

C. Campus Deployments of Autonomous Buses

University campuses serve as effective testbeds for AVs due to their controlled yet dynamic environments (see Figure 1). Studies have shown that repeated exposure influences pedestrian trust and acceptance. The University of Michigan's Mcity shuttle program demonstrated that campus environments offer unique advantages for studying AV-pedestrian interactions due to their mix of regular users and first-time encounters [28]. Similarly, the long-term presence of autonomous buses on the Campus of Linköping University in Sweden made it possible to examine how long-term exposure to AVs affected cyclists. They found that cyclists with longer exposure reported more negative attitudes due to unexpected behaviors like sudden braking [29]. Further research in this area highlights the complex social dimension of traffic, noting that autonomous vehicles must learn to comply with the expectations of human road users, such as cyclists, to ensure smooth and socially acceptable coordination in less structured environments [6]. The Singapore-MIT Alliance for Research and Technology (SMART) focused on AVs navigating pedestrian crossings and intersections, aiming to raise public awareness and acceptance [30].

These studies highlight key insights: repeated AV exposure influences trust and attitudes, people's diversity affects interaction patterns, and effective communication systems are crucial for conveying AV intentions.

D. Social Status and Representation of AVs

Research on AV-pedestrian interactions has highlighted the importance of social cues and attributions in shaping pedestrian behavior. Previous studies show that pedestrians' willingness to yield is influenced not only by traffic rules and safety considerations but also by the social meanings attached to AVs [31], [32]. Research has investigated how the perceived representation of AVs, whether as the passengers it transports or as an autonomous system, shapes social responses. For example, studies in human-robot interaction suggest that people attribute more social status and priority to robots when they are perceived as acting on behalf of humans, rather than as autonomous agents ('just a robot, it can wait') [33]. Concerning shuttle buses, it was found that when the buses run empty, they are not perceived as serving their purpose and are treated less cautiously by cyclists (e.g., close overtaking provoking harsh braking) [29]. Taken together, this suggests that both social status and representation could play a role in how pedestrians decide whether to yield to AVs.

E. Research Questions

While previous studies have explored AV-pedestrian interactions mostly with cars, our study offers insights into pedestrian behavior from deploying an autonomous shuttle bus with two communication modalities (audio and visual) in a real-world setting, a shared space on a university campus. In addition to behavioral observation, we gathered subjective measures on pedestrians' perceptions of the AV to investigate the following research questions:

- **RQ1**: Does prior experience with AVs influence pedestrian trust in their safety?
- **RQ2**: How does the perceived social status and representation of AVs affect pedestrian willingness to yield?
- **RQ3**: What attitudes do students on campus report to sharing their space with the AV?

Fig. 2. Real-world experimental setting showing the AutoBus with the ’DRIVING’ visual cue displayed on its front LED screen (left), and a top-view snapshot of the 480-meter experimental route on campus (right).

III. METHOD

A. *Experimental Scenario*

1) Outdoor Setting: The experiment was conducted on the campus of RPTU Kaiserslautern, a technical university in Germany with a diverse and international student population of approximately 20.000. The campus features numerous pedestrian-only zones and shared paths, making it a suitable environment for studying pedestrian-AV interactions. On campus, the AutoBus, an autonomous shuttle bus, navigated pedestrian zones while adhering to the Ethical Rules for Automated and Connected Vehicular Traffic set by the German Ethics Commission in 2017[2]. The AutoBus has been present on campus for approximately four years as part of a long-term research project, operating intermittently for various tests. This long-term, sporadic presence ensures a mix of participants: those who were already familiar with the vehicle and those encountering it for the first time. The study was conducted over four randomized days in January 2025. To capture interactions during a period of high pedestrian traffic, the bus operated for several intervals within the university’s peak lunchtime window (11:00 - 13:00) on each of the four days, circling a 480m path. A male safety driver was present inside the bus at all times to intervene in case of emergencies. Figure 2 illustrates the real-world experimental setting, with the driven path shown in Figure 2 (right). Four experimenters were stationed along this route (indicated by yellow markers) and approached pedestrians immediately after they encountered the bus. These participants were invited to complete the survey via a QR code on their mobile devices, allowing them to respond either immediately to minimize memory decay or later at their convenience.

2) Autonomous Bus Operation: The AutoBus maintained a constant speed of 4 km/h when unobstructed, communicating its status via two modalities. The messages included: "PARKING", "DRIVING", "STARTING TO DRIVE", and "PLEASE GIVE WAY!". A demo can be viewed here: https://youtu.be/DnPIw3_P7Ao.

The decision-making process for triggering communication cues followed a hierarchical logic: visual cues (LED display) remained continuously active to provide ambient awareness for all pedestrians. Auditory cues were triggered conditionally using the following algorithm:

Step 1: **Pedestrian Detection**: When the perception system detected a pedestrian within 10 meters
Step 2: **Awareness Assessment**: The skeleton tracking system classified the pedestrian’s awareness state based on head orientation and body posture
Step 3: **Cue Activation**: If classified as “unaware,” an auditory alert was automatically triggered
Step 4: **Escalation Protocol**: If the pedestrian remained stationary and did not yield within approximately 15 seconds after the initial auditory warning, the system issued a single short horn signal
Step 5: **Safety Override**: The safety driver could intervene at any point if necessary

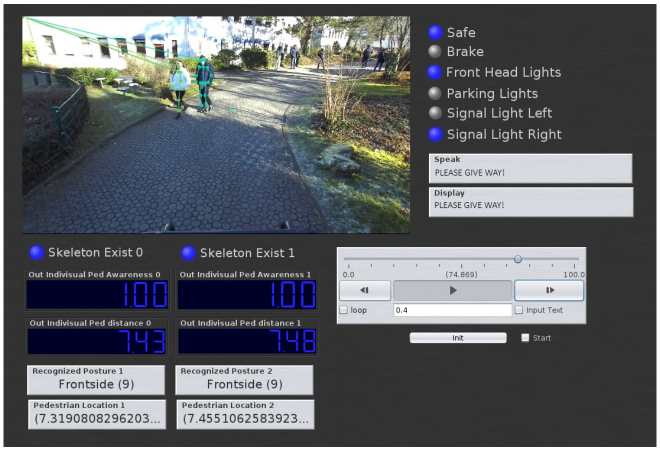
This adaptive triggering approach minimized noise pollution while providing targeted alerts to pedestrians who appeared unaware of the bus’s presence.

- **Visual Cue**: A custom-made, weatherproof (IP65) LED display (420x84 px, 6,000 cd/m²) provided a continuously active visual signal. It displayed the vehicle’s real-time status, as mentioned above, in bright green text to ensure high legibility in various lighting conditions.
- **Auditory Cue**: An external speaker system delivered auditory alerts. Using a synthesized male, English-language voice, the system verbally announced the same message currently shown on the LED display. Unlike the continuous visual cue, these alerts were triggered automatically only when the perception system detected a pedestrian within a 10-meter proximity who was classified as unaware of the bus.

3) Pedestrian Detection System: The AutoBus utilized an onboard camera and real-time skeleton tracking to detect and prioritize the two closest pedestrians, collecting observational

---

[2] `https://www.bmv.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.html`



● Safe
○ Brake
● Front Head Lights
○ Parking Lights
○ Signal Light Left
● Signal Light Right

Speak
PLEASE GIVE WAY!

Display
PLEASE GIVE WAY!

● Skeleton Exist 0
● Skeleton Exist 1

Out Indisivual Ped Awareness 0
1.00

Out Indisivual Ped Awareness 1
1.00

Out Indisivual Ped distance 0
7.43

Out Indisivual Ped distance 1
7.48

Recognized Posture 1
Frontside (9)

Recognized Posture 2
Frontside (9)

Pedestrian Location 1
(7.3190808296203...

Pedestrian Location 2
(7.4551062583923...

☐ loop 0.4 ☐ Input Text

Fig. 3. Internal GUI of AutoBus.

data over a total recording duration of 2 hours and 17 minutes. A graphical user interface (GUI) inside the bus (see Figure 3) displayed the active command (text and/or voice), current safety status, detected pedestrian skeletons with their relative positions, and estimated awareness states (e.g., looking at the bus, distracted by a phone).

4) Participants: Pedestrians who walked toward and crossed the AutoBus, being detected by the aforementioned pedestrian detection system ($N = 256$), were invited to participate in a survey via QR codes following their interaction. Of these, $n = 67$ responded. Responses were deemed valid if the participant completed all questions related to the core research questions, resulting in $n = 58$ valid responses (47 male, 11 female) being retained for analysis to supplement the recorded behavioral labels. The mean age was 25 years (SD = 4.58), ranging from 18 to 42 years. Participants represented a range of academic disciplines at the university, with the largest groups from Computer Science ($n = 16$), Biology ($n = 13$), and Electrical and Computer Engineering ($n = 11$). All participants provided informed consent and completed the full procedure. No data was excluded.

B. Measures

1) Observational: Pedestrian awareness and behavior were analyzed using real-time skeleton tracking [13][Section 4]. The system's estimation of awareness, based on posture and head orientation, was validated against a manually coded ground truth from 50 interaction clips, achieving an accuracy of 89%. At the same time, facial expressions were subjectively assessed by the first two authors (1 male, 1 female) who were trained on a standardized coding scheme. Although an automated facial expression recognition system is integrated into the AutoBus platform, manual annotation was employed in the present study due to several methodological considerations: (1) the uncontrolled outdoor environment presented challenges including variable lighting conditions, occlusions from masks or accessories, and varying distances between pedestrians and cameras, which reduces the reliability of automated systems; (2) manual coding allowed for nuanced interpretation of contextual factors (e.g., distinguishing bus-related expressions from social interactions); and (3) our focus on broader behavioral patterns rather than frame-by-frame expressions made human judgment more appropriate for this exploratory study. Future work incorporating real-time automated expression recognition could enable larger-scale analysis and adaptive eHMI responses. To ensure reliability, the coders independently labeled a subset of 20% of the interactions. Inter-rater reliability was calculated using Cohen's Kappa, which showed substantial agreement ($\varkappa$ = 0.82). Any disagreements on the remaining data were resolved through discussion.

To categorize pedestrian interactions with the AutoBus, behaviors were coded into predefined categories. These categories were developed through a grounded theory approach, where two researchers independently reviewed 25% of the video footage to identify recurring behaviors. The initial codes were then discussed, refined, and consolidated into the final set of categories:

- **Interacted with the bus**: Direct engagement detected.
- **Approaching**: Individuals walking toward the bus, possibly intending to cross.
- **Looked at bus**: Head movement indicating attention toward the bus or eHMI.
- **Navigated around**: Adjustments in path due to the bus obstructing the route.
- **Awareness shift (Text/Voice)**: Pedestrians who became aware via text display or auditory cue.
- **Moving away**: Pedestrians who changed direction after noticing the bus.
- **Paused to observe**: Pedestrians who stopped to observe, sometimes taking photos/videos.

2) Post-Experiment Survey: The post-experiment survey assessed participants' prior experience with the AutoBus, attitudes toward its presence on campus, trust in its safe operation, and willingness to yield based on perceived social status and representation [29], [33]. Overall trust was rated on a 5-point Likert scale (1 = Not at all, 5 = Completely), and relative safety was categorized as less safe, equally safe, or safer than a human driver [29]. Perceived social status was measured by asking participants to rate the bus relative to themselves, and representation was assessed by whether the bus was perceived as acting autonomously or on behalf of a human [33]. All survey items and response scales are reported in Appendix A.

TABLE I
CONTINGENCY TABLE: AUDIO VS. VISUAL

Noticed Textual Cue	Noticed Audio Cue No	 Yes	Total
No	14	14	28
Yes	18	12	30
Total	32	26	58

TABLE II

OBSERVED PEDESTRIAN BEHAVIOR

Category	**Count**
Interacted with the bus	256
Approaching	223
Looked at bus	164
Navigated around	107
Unaware → Aware (Voice)	68
Unaware → Aware (No Voice)	47
Moving away	33
Paused to observe	29

Facial Expression	**Count**
Neutral	135
Smiled	86
Unrecognized	35

IV. RESULTS & ANALYSIS

Data analyses were conducted using JASP (v0.95.2). Due to non-normal distributions in the data (Shapiro-Wilk, $p < .001$), non-parametric statistical tests were applied throughout the analysis.

A. *Manipulation Checks*

1) Prior Experience: This analysis served as a manipulation check to assess whether prior exposure to AutoBus influenced participants' recognition of its communicative features. A Mann-Whitney U test revealed a significant increase in visual cue recognition ($U = 255.000, p = .002, r_{rb} =$ -0.387). However, no significant difference was found for auditory information ($U = 348.000, p = .110, r_{rb} =$ -0.163), indicating that prior exposure did not change audio perception (see Table I).

2) Cue Perception: Observational data from the campus driving experiments (see Table II) showed that 28 participants did not notice visual cues, equally split between those who noticed ($n = 14$) and did not notice ($n = 14$) audio cues. Conversely, among the 30 participants who noticed visual cues, 18 failed to perceive the auditory cue, suggesting that visual cues were more consistently recognized than auditory ones.

B. *Observed Behavior*

Pedestrians frequently looked at the bus ($n = 164$) or paused to observe ($n = 29$), reinforcing the stronger impact of visual cues. However, auditory cues seemed like a suitable modality to create awareness of the AV's presence as more individuals became aware of the bus when voice cues were present ($n = 68$) compared to when absent ($n = 47$). Additionally, most participants displayed neutral or positive (smiling) facial expressions as reported in Table II, potentially indicating interest or acceptance of the AutoBus.

C. *RQ1: Trust in AVs Safe Operation*

A Chi-square test showed no significant association between prior experience with the AutoBus and pedestrians' relative safety judgments ($\chi^2 = 3.091, p = .213$). An independent-samples Mann-Whitney U test likewise indicated no significant difference in trust in automated bus driving between participants with and without prior experience ($U = 435.5, p = .631, r_{rb} = -0.047$).

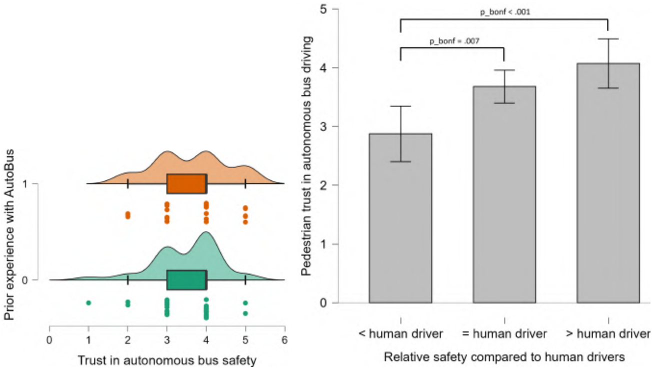


Fig. 4. A combined scatterplot and boxplot of the interplay between AutoBus pre-experience and overall trust in the safer operation of the buses (green = no pre-experience; orange = pre-experienced participants) *(left)*. Box plot indicating perceived safety of autonomous buses from a pedestrian perspective relative to a human driver *(right)*.

A Kruskal-Wallis test revealed a significant effect of perceived relative safety on trust ($H = 15.50, p < .001$), with a moderate-to-strong effect size ($\epsilon^2 = .272, 95\%\,CI\,[0.101, 0.508]$). Dunn's post hoc tests with Bonferroni correction showed significantly lower trust among participants who rated the bus as less safe than a human driver compared to those rating it as equally safe ($p_{bonf} = .007$) or safer ($p_{bonf} < .001$), while no difference was found between the latter two groups ($p_{bonf} = .580$) (see Figure 4(right)).

Descriptively, 16 participants rated the autonomous bus as less safe, 28 as equally safe, and 14 as safer than a human driver.

D. *RQ2: Willingness to Yield*

Spearman's correlation revealed a moderate positive association between perceived social status of the AutoBus and willingness to yield in competitive scenarios ($\rho = .460, p < .001$). No significant correlations were found between willingness to yield and perceptions of the bus as representing AI versus a human operator ($p > .05$). Other correlations between social status and yielding were not significant.

Wilcoxon signed-rank tests showed no significant differences in willingness to yield between priority and competition conditions across scenarios ($p > .05$). Overall, these results indicate that pedestrians' yielding behavior is more strongly related to the perceived social status of the AutoBus than to whether it is seen as acting autonomously or on behalf of a human. Descriptive statistics are shown in Figure 5.

TABLE III

DESCRIPTIVE STATISTICS: TRUST IN AVS ON CAMPUS

		95% CI Mean		
	Mean	Upper	Lower	Std. Dev
Q10. AV campus presence	4.000	4.246	3.754	0.937
Q11. AV campus driving	3.845	4.109	3.581	1.005
Q12. Overall AV impact	3.914	4.157	3.671	0.923

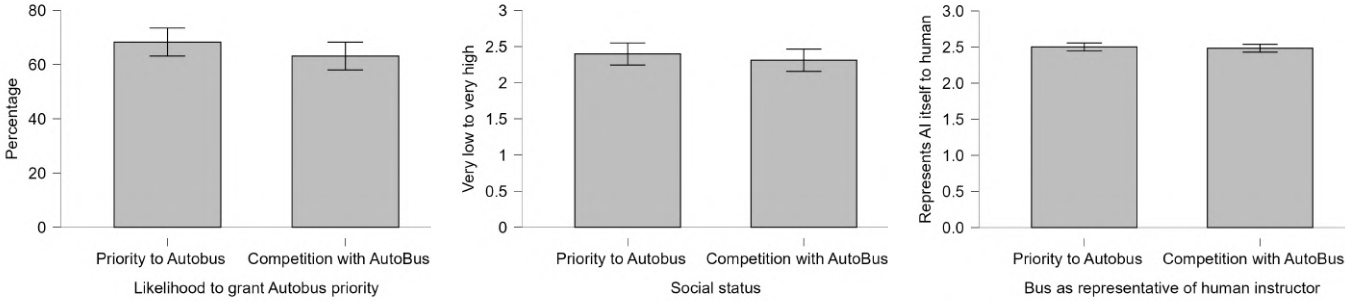


Fig. 5. *(left)*: Likelihood of granting priority to the AutoBus. *(middle)*: Perceived social status of the AutoBus compared to the pedestrian. *(right)*: Perception of the AutoBus as representing AI itself or a human instructor.

E. *RQ3: AVs on Campus*

Descriptive analyses summarize students' attitudes toward the presence of the AutoBus on campus, its operation along campus paths, and its overall impact.

Mann-Whitney U showed no significant differences in perceptions of AV presence between those with and without prior experience (Q10: U = 355.000, p = .159; Q11: U = 391.000, p = .345; Q12: U = 379.500, p = .277). Effect sizes (rank-biserial correlation were small (Q10: $r_{rb} = -0.147$, Q11: $r_{rb} = -0.060$, Q12: $r_{rb} = -0.088$), suggesting that prior experience had little to no impact on AV perceptions. Descriptive statistics indicated that overall attitudes were neutral to positive (Means: Q10 = 4.00, Q11 = 3.85, Q12 = 3.91). These findings inform us of pedestrians' perceptions of AVs, as shown in Table III.

V. DISCUSSION

This study examined pedestrian responses to an autonomous bus (AutoBus) operating in a university campus environment. The results contribute to understanding how trust, social perceptions, and prior experience influence AV-pedestrian interactions in public spaces.

A. *Discussion of Results*

With regard to **RQ1**, trust in AVs safe operation did not significantly differ between participants with and without prior exposure to the AutoBus. However, perceived safety relative to human drivers showed a meaningful effect: participants who believed AVs were less safe than human drivers expressed significantly lower trust. This suggests that trust in AVs is shaped more by subjective safety perceptions than by familiarity or experience alone. Unlike findings from Thellman et al. [29], where long-term exposure led to more negative attitudes toward AVs, our participants did not have this long-term exposure to the AutoBus, so future studies are needed to test long-term effects on trust.

For **RQ2**, willingness to yield was positively associated with the perceived social status of the AV, especially when pedestrians imagined a scenario where they had to compete for space with the AV. However, no correlation was found between willingness to yield and the perceived representation of the AV as transporting passengers or merely as an autonomous robot. These findings indicate that social status attribution to AVs, rather than their perceived representation, play a stronger role in priority negotiations in AV-pedestrian conflicts. This aligns with prior work [33], which showed that social status attributions to AVs significantly shape how individuals negotiate shared space use with AVs.

In addressing **RQ3**, participants' general attitudes toward AV presence on campus were largely neutral to positive. No significant differences were found between experienced and first-time observers across questions regarding AV presence and behavior, suggesting that passive exposure to the AutoBus did not meaningfully alter perceptions. The prevalence of neutral or smiling facial expressions further indicates general acceptance of the AV. As observed in other university studies, such as in Mcity and SMART [28], [30], campus settings offer a viable environment for capturing naturalistic AV-human interactions across varied familiarity levels.

One caveat to consider with real-world studies is that stimuli presentation cannot be as standardized as in lab studies due to, for example, noise or glare. Hence, it needs to be discussed how the differential perception of the communication cues might have influenced our main findings. The fact that visual cues were more consistently recognized than the auditory cues suggests that participants' trust and yielding behaviors were likely shaped more by what they saw on the bus's display and its physical movement than by its auditory alerts. While this study reflects the reality of applying eHMIs in the real world, this has to be taken into account when interpreting the results. The auditory cue was delivered using a male synthetic voice (Acapela "Graham"), and the safety operator present in the vehicle was male and seated in the conventional driver position. Although not experimentally manipulated, these gendered cues may have implicitly shaped perceptions of authority or competence in the interaction context.

Overall, the findings support earlier work [21], [22] emphasizing the role of multimodal communication in shaping pedestrians' behavior to AVs and extend them to autonomous shuttle buses. While prior exposure alone did not significantly influence trust or behavior, perceptions of safety and status strongly influenced responses. These findings suggest that future AV design should not only focus on communication modalities but also on social perceptions of AVs to enhance cooperative behavior.

B. Limitations and Future Work

Several limitations should be noted. First, facial expressions were coded manually after the experiment, and some recorded expressions may have been unrelated to the bus interaction due to surrounding social context. This approach was chosen over automated methods due to the challenges of accurately classifying nuanced expressions in uncontrolled outdoor lighting and at variable distances. Second, the awareness detection algorithm occasionally misclassified pedestrians who had previously noticed the bus but later looked away. Additionally, reliance on front-facing cameras may have excluded individuals who interacted with the bus from lateral or rear positions. Several factors may have limited the auditory cue's effectiveness. Its design differed from the continuously active visual cue, as it was triggered only upon pedestrian detection (a choice made to reduce noise pollution). Furthermore, its impact could have been affected by uncontrolled variables on our international campus, such as participants' English proficiency or use of headphones. Because the study was conducted in a multicultural campus environment, cultural associations between gender and driving roles may have influenced interpretations of the male voice cue and visible safety driver. Future work should systematically examine how voice gender and operator presence affect pedestrian responses across contexts.

The sample, although representative of a university setting, limits the generalizability of findings to broader urban populations. Our study focused exclusively on pedestrians. As noted by Thellman et al. [29], cyclists' long-term interactions with autonomous buses can lead to more negative attitudes. Hence, future work is needed on how other road users, such as bikers and scooter riders, interact with and perceive the AutoBus. Finally, future work will explore implementing state-of-the-art open-weight models available on the Ollama framework, such as LLaMA 3.2 Vision and LLaVA, to extract contextual insights from camera sensors as well as assess pedestrian perceptions before and after riding as passengers in the AutoBus.

VI. CONCLUSION

This exploratory study provides preliminary insights into real-world pedestrian interactions with an autonomous shuttle bus in a campus setting. Results highlight that perceptions of social status and AV safety have a stronger influence on pedestrian behavior than prior experience or the perceptions of the AV's passenger representation. These findings provide implications for the design of AVs not only based on communication modality but also on social perceptions. As AV deployment expands, continued research in varied environments and with larger populations will be essential to support inclusive, trustworthy mobility systems.

ACKNOWLEDGEMENT

During manuscript preparation, the authors used ChatGPT (OpenAI, GPT-5.2) solely for language polishing. The authors reviewed and edited all text and accepted full responsibility for the manuscript.

APPENDIX

APPENDIX A. QUESTIONNAIRE ITEMS

Item No.	Question Item and Response Scale
1.	Consent Statement
Part 1	**Prior Experience and Cue Perception**
2.	Do you have prior experience with AutoBus? *[Yes/No]*
3.	If answered Yes above, how? *[Online / In-person at RRLab / On the campus]*
4.	Do you have prior experience with being a passenger of the AutoBus? *[Yes/No]*
5.	Did you notice the textual information displayed on the AutoBus? *[Yes/No]*
6.	Did you hear and understand the audio information provided by the AutoBus? *[Yes/No]*
7.	Do you have prior experience with another autonomous shuttle? *[Yes/No]*
8.	If answered Yes above, which one(s)? *[Open-ended]*
9.	Do you have prior experience with being a passenger of the other autonomous shuttle? *[Yes/No]*
Part 2	**Attitudes Towards AutoBus**
10.	How do you feel about the presence of the autonomous buses on campus? *5-point scale: 1 (do not like them at all) to 5 (like them a lot)*
11.	How do you feel about the autonomous buses driving on the campus path? *5-point scale: 1 (very bad) to 5 (very good)*
12.	Do you think the overall presence of the autonomous buses on campus is good or bad? *5-point scale: 1 (very bad) to 5 (very good)*
13.	Because... *[Open-ended]*
Part 3	**Trust**
14.	Do you trust the autonomous buses to drive safely? *5-point scale: 1 (not at all) to 5 (completely)*
15.	As a pedestrian, do you trust the self-driving buses to drive safely? *[More unsafe than a human driver / Equally safe / Safer than a human driver]*
16.	Please report any potentially dangerous encounters and provide possible solutions. *[Open-ended]*
Part 4	**Yielding Scenario (Bus has Priority)**
17.	How likely is it that you would grant the bus priority? *[Range: 0–100]*
18.	What social status would you assign to the bus in comparison to yourself? *2-point scale: 1 (Very low) to 2 (Very high)*
19.	To what extent do you see the bus as a representative of the human who instructed it? *5-point scale: 1 (Represents AI itself) to 5 (Represents human)*
20.	What did you base your decision on? *[Open-ended]*
Part 5	**Yielding Scenario (Competition for Space)**
21.	How likely is it that you would grant the bus priority? *[Range: 0–100]*
22.	What social status would you assign to the bus in comparison to yourself? *2-point scale: 1 (Very low) to 2 (Very high)*
23.	To what extent do you see the bus as a representative of the human who instructed it? *5-point scale: 1 (Represents AI itself) to 5 (Represents human)*
24.	What did you base your decision on? *[Open-ended]*
Part 6	**Final Comments**
25.	Do you want to comment on the study or the bus? *[Open-ended]*

REFERENCES

[1] H. Du, G. Zhu, and J. Zheng, "Why travelers trust and accept self-driving cars: An empirical study," *Travel behaviour and society*, vol. 22, pp. 1–9, 2021.

[2] A. R. Palmeiro, S. van der Kint, L. Vissers, H. Farah, J. C. de Winter, and M. Hagenzieker, "Interaction between pedestrians and automated vehicles: A wizard of oz experiment," *Transportation research part F: traffic psychology and behaviour*, vol. 58, pp. 1005–1020, 2018.

[3] M. A. Richter, M. Hagenmaier, O. Bandte, V. Parida, and J. Wincent, "Smart cities, urban mobility and autonomous vehicles: How different cities needs different sustainable investment strategies," *Technological Forecasting and Social Change*, vol. 184, p. 121857, 2022.

[4] M. Prédhumeau, A. Spalanzani, and J. Dugdale, "Pedestrian behavior in shared spaces with autonomous vehicles: An integrated framework and review," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 438–457, 2021.

[5] B. Färber, "Communication and communication problems between autonomous vehicles and human drivers," in *Autonomous driving: Technical, legal and social aspects*. Springer, 2016, pp. 125–144.

[6] H. R. Pelikan, "Why autonomous driving is so hard: The social dimension of traffic," in *Companion of the 2021 ACM/IEEE international conference on human-robot interaction*, 2021, pp. 81–85.

[7] H. R. Pelikan, B. Mutlu, and S. Reeves, "Making sense of public space for robot design," in *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2025, pp. 152–162.

[8] G. Markkula, R. Madigan, D. Nathanael, E. Portouli, Y. M. Lee, A. Dietrich, J. Billington, A. Schieben, and N. Merat, "Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic," *Theoretical Issues in Ergonomics Science*, vol. 21, no. 6, pp. 728–752, 2020.

[9] C. Ackermann, M. Beggiato, L.-F. Bluhm, A. Löw, and J. F. Krems, "Deceleration parameters and their applicability as informal communication signal between pedestrians and automated vehicles," *Transportation research part F: traffic psychology and behaviour*, vol. 62, pp. 757–768, 2019.

[10] M. Prédhumeau, L. Mancheva, J. Dugdale, and A. Spalanzani, "Agent-based modeling for predicting pedestrian trajectories around an autonomous vehicle," *Journal of Artificial Intelligence Research*, vol. 73, pp. 1385–1433, 2022.

[11] I. Nastjuk, B. Herrenkind, M. Marrone, A. B. Brendel, and L. M. Kolbe, "What drives the acceptance of autonomous driving? an investigation of acceptance factors from an end-user's perspective," *Technological Forecasting and Social Change*, vol. 161, p. 120319, 2020.

[12] K. Othman, "Public acceptance and perception of autonomous vehicles: a comprehensive review," *AI and Ethics*, vol. 1, pp. 355–387, 2021.

[13] Q. H. Jan, "Towards a smart interaction strategy in driverless minibuses for optimizing pedestrian interaction," Ph.D. dissertation, Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, 2024.

[14] J. She, J. Neuhoff, and Q. Yuan, "Shaping pedestrians' trust in autonomous vehicles: an effect of communication style, speed information, and adaptive strategy," *Journal of Mechanical Design*, vol. 143, no. 9, p. 091401, 2021.

[15] D. Dey, A. Matviienko, M. Berger, B. Pfleging, M. Martens, and J. Terken, "Communicating the intention of an automated vehicle to pedestrians: The contributions of ehmi and vehicle behavior," *it-Information Technology*, vol. 63, no. 2, pp. 123–141, 2021.

[16] N. Lownes, S. Rezwana, M. R. R. Shaon, E. Jackson, *et al.*, "Pedestrian behavior and interaction with autonomous vehicles (phase ii)," 2023.

[17] D. Dey, A. Habibovic, A. Löcken, P. Wintersberger, B. Pfleging, A. Riener, M. Martens, and J. Terken, "Taming the ehmi jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces," *Transportation Research Interdisciplinary Perspectives*, vol. 7, p. 100174, 2020.

[18] M. Bied, B. Bruno, and A. Vinel, "Autonomous vehicles as social agents: Vehicle to pedestrian communication from v2x, ehmi and hri perspectives," in *2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2024, pp. 86–91.

[19] K. Kühnlenz and B. Kühnlenz, "Autonomous vehicle ehmi communication modalities: pedestrian preferences in crossing scenarios," *at-Automatisierungstechnik*, vol. 72, no. 12, pp. 1160–1168, 2024.

[20] H. Sahin, K. Daudrich, H. Müller, and S. C. J. Boll, "Signaling yielding intent with ehmis: The timing determines an efficient crossing," in *Adjunct Proceedings of the 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2021)*, 2021, pp. 5–9.

[21] C. Ackermann, M. Beggiato, S. Schubert, and J. F. Krems, "An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles?" *Applied ergonomics*, vol. 75, pp. 272–282, 2019.

[22] S. Chng, M. Kawasumi, M. Al-Ahda, C. Mangla, and L. Cheah, "Evaluating the use of graphic symbols in autonomous vehicle-to-pedestrian communication," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 389–394.

[23] K. De Clercq, A. Dietrich, J. P. Núñez Velasco, J. De Winter, and R. Happee, "External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions," *Human factors*, vol. 61, no. 8, pp. 1353–1370, 2019.

[24] H. R. Pelikan and M. F. Jung, "Designing robot sound-in-interaction: The case of autonomous public transport shuttle buses," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 172–182.

[25] S. Deb, D. W. Carruth, and C. R. Hudson, "How communicating features can help pedestrian safety in the presence of self-driving vehicles: virtual reality experiment," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 2, pp. 176–186, 2020.

[26] T. Nalini, S. S. Velan, M. Shanmuganathan, K. Kajendran, and P. Radhakrishnan, "Audible pedestrian warning system using embedded system," *Measurement: Sensors*, vol. 25, p. 100635, 2023.

[27] P. Bazilinskyy and J. de Winter, "Auditory interfaces in automated driving: an international survey," *PeerJ Computer Science*, vol. 1, p. e13, 2015.

[28] M. M. Nesheli, L. Li, M. Palm, and A. Shalaby, "Driverless shuttle pilots: Lessons for automated transit technology deployment," *Case studies on transport policy*, vol. 9, no. 2, pp. 723–742, 2021.

[29] S. Thellman, F. Babel, and T. Ziemke, "Cycling with robots: How long-term interaction experience with automated shuttle buses shapes cyclist attitudes," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 1043–1047.

[30] H. Eng, "Embracing the future of land transportation: valuing flexibility in design and technology options for autonomous vehicle developments in singapore," Ph.D. dissertation, Massachusetts Institute of Technology, 2017.

[31] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE transactions on intelligent transportation systems*, vol. 21, no. 3, pp. 900–918, 2019.

[32] L. Fridman, B. Mehler, L. Xia, Y. Yang, L. Y. Facusse, and B. Reimer, "To walk or not to walk: Crowdsourced assessment of external vehicle-to-pedestrian displays," *arXiv preprint arXiv:1707.02698*, 2017.

[33] F. Babel, P. Hock, K. Winkle, I. Torre, and T. Ziemke, "The human behind the robot: Rethinking the low social status of service robots," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 1–10.