Prosociality in Microtransit*

Divya Sundaresan
NC State University
Raleigh, United States
dsundar3@ncsu.edu

Akhira Watson
NC State University
Raleigh, United States
awatson9@ncsu.edu

Eleni Bardaka
NC State University
Raleigh, United States
ebardak@ncsu.edu

Crystal Chen Lee
NC State University
Raleigh, United States
cchen32@ncsu.edu

Christopher B. Mayhorn
NC State University
Raleigh, United States
cbmayhor@ncsu.edu

Munindar P. Singh
NC State University
Raleigh, United States
mpsingh@ncsu.edu

ABSTRACT

We study *(public) microtransit*, a type of transportation service wherein a municipality offers point-to-point rides to residents, for a fixed, nominal fare. Microtransit exemplifies practical resource allocation problems that are often over-constrained in that not all ride requests (pickup and dropoff locations at specified times) can be satisfied or satisfied only by violating soft goals such as sustainability, and where economic signals (e.g., surge pricing) are not applicable—they would lead to unethical outcomes by effectively coercing poor people. *Prosociality* refers to an attitude or behavior that is intended to benefit others. This paper demonstrates a computational approach to prosociality in the context of a *(public) microtransit* service for disadvantaged riders.

This paper describes an interdisciplinary study of prosociality in microtransit between a transportation researcher, psychologists, a social scientist, and AI researchers. Our contributions are these: (1) empirical support for the viability of prosociality in microtransit (and constraints on it) through interviews with drivers and focus groups of riders; (2) a prototype mobile app demonstrating how our prosocial intervention can be combined with the transportation backend; (3) a reinforcement learning approach to model a rider and determine the best interventions to persuade that rider toward prosociality; and (4) a cognitive model of rider personas to enable evaluation of alternative interventions.

KEYWORDS

computational social systems, multiagent systems, cognitive modeling, reinforcement learning

ACM Reference Format:
Divya Sundaresan, Akhira Watson, Eleni Bardaka, Crystal Chen Lee, Christopher B. Mayhorn, and Munindar P. Singh. 2025. Prosociality in Microtransit. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

*The full version of this paper appears in the Journal of Artificial Intelligence Research [25].

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

Transportation is essential for residents to go to work, obtain healthcare, shop for food, or otherwise engage in civic life. *(Public) microtransit* services are a kind of public transit system wherein a municipality offers point-to-point rides to residents [20]. Microtransit can involve a mixture of on-demand and commuter programs. However, current microtransit approaches face challenges in that resources (minivans and drivers) are expensive and, at the same time underutilized, while many ride requests are declined for lack of availability. Importantly, microtransit cannot ethically rely on price signals to manage high demand since the target user population is vulnerable and economically disadvantaged.

We consider a sociotechnical system (STS) [22] as a multistakeholder cyberphysical system. An STS has a social tier of people and organizations and a technical tier of cyberphysical resources and data. In the present setting, a municipal microtransit service constitutes an STS. Its stakeholders (including users and providers, i.e., riders, drivers, and the city transit authority) form the social tier of the STS. Its cyberphysical resources and data (i.e., vehicles and the associated information technology to request rides) form the technical tier of the STS. We posit that problems that may be difficult to solve at the technical tier can be made tractable through interventions in the social tier.

We describe an approach to improving the efficiency and effectiveness of microtransit by promoting prosocial attitudes among riders, reflected in their adjusting their preferences to facilitate sharing rides [5]. We adopt a public-centric approach wherein we view riders as agents at the center of the multiagent microtransit system to ensure that the system is accepted by riders as safe, reliable, and trustworthy [24]. To build an agent that suggests acceptable interventions, our first challenge is to understand riders and tailor suggestions accordingly for maximum effect. Some people are willing to adjust their departure or arrival times but not their pickup and dropoff locations (temporal versus spatial flexibility). Similarly, riders may vary in their willingness to adjust their trips based on who stands to benefit. The foregoing motivation leads us to examine the following research questions.

$RQ_{tolerance}$ Can we learn riders' spatial tolerances to suggest optimal spatial adjustments?

$RQ_{empathy}$ Can we learn riders' empathetic tendencies to persuade them to adjust?

$RQ_{profile}$ Could considering rider profile data lead to a better (non-naive) starting point?

1.1 Approach and Contributions

To understand riders, we apply established social science methods to elicit their needs (requirements, risk attitudes, and values). The requirements refer to preferences regarding their trips and their flexibility. The risk attitudes refer to their views of risks such as being late or walking in the dark. The values refer to empathy and helping others in need. As our computational method, we apply machine learning to learn to persuade riders to relax their requests in light of their risk attitudes and values. For the purposes of evaluation, we capture rider needs and responsiveness to various persuasive messages.

Figure 1 illustrates our setting and the plan of the paper. As envisioned, a fielded solution would involve riders (humans) engaging with the CARS agent through a mobile app. The agent would model riders and attempt to persuade them to be prosocial (as described below) to improve overall system performance and rider satisfaction. In our experiments, we use a simplified version of the agent (considering a simplified environment) along with simulated riders, and address the research questions $\mathrm{RQ}_{\mathrm{tolerance}}$, $\mathrm{RQ}_{\mathrm{empathy}}$, and $\mathrm{RQ}_{\mathrm{profile}}$ introduced in Section 1. The riders are simulated based on the ACT-R cognitive architecture, including some parameters based on the Social Value Orientation (SVO) literature.

1.2 Novelty

Previous research on AI for transportation and urban mobility has not tackled the challenges we address here. Some of it accommodates only predetermined rider preferences, such as driver competence and vehicle safety [19, 26]. Other research focuses on economic incentives [9]. We consider users (i.e., riders) as central to the system and eliminate economic incentives in favor of persuasion toward prosociality. Some researchers have considered ways to enhance the attractiveness of alternative options to users, such as goal setting, personalized messaging, social comparison, and gamification [1, 10]. We build on their ideas, but we expand persuasive messaging to accommodate considerations of empathy. In addition, we consider an adaptive approach to persuade riders while learning their preferences, risk attitudes, and values—and respecting those risk attitudes and values in the persuasions attempted.

2 BACKGROUND ON MICROTRANSIT

We consider the setting of *(public) microtransit* services that are emerging in rural areas in the US. People with disabilities or those who are elderly or poor must rely on public transportation. Rural areas have a low population density and fixed-route transit services (such as bus and rail) prove unviable since they are both expensive and underutilized. As a result, municipalities such as Wilson, North Carolina (our partner in this study) have shut down their fixed route transit and replaced it with microtransit through a small fleet of minivans, each able to hold a driver and up to six passengers. Wilson, with a population of 40,000, was the first in North Carolina to implement a city-wide microtransit system, called RIDE, operated by Via (RIDE's service provider).

During a workshop we conducted with the key stakeholders of RIDE, we learned that unfortunately, during the morning and afternoon peak periods, a substantial fraction of ride requests received cannot be served. This is a major problem because, based on a survey conducted by Via, about 60% of the riders in Wilson use microtransit mainly for work and medical appointments. In addition, 86% are carless and 57% earn less than $25K per year. Hence, many riders face daily struggles with the microtransit but cannot switch to other modes due to a lack of alternative travel options. Despite the high demand, the microtransit vans (which can fit up to six passengers) remain highly underutilized.

3 UNDERSTANDING STAKEHOLDER NEEDS

We conducted interviews with all the key stakeholders. One group consists of the operational transportation managers in Wilson, from whom we learned about the economic constraints on the service. From a second group, drivers, we learned about their estimates of rider flexibility.

We conducted focus groups in Wilson to understand the largest and most important group of stakeholders in our setting: riders. Unlike one-on-one interviews, focus groups reveal the similarities and differences between participants in a social setting [15]. Semistructured focus groups are particularly useful for studies of how people make sense of a particular phenomenon or experience [12, 13]. 165 microtransit riders signed up for the five focus group sessions we organized (eIRB# 25553). We conducted five focus group sessions for a total of 32 participants, selected at random from the 165 candidates. We invited six or seven participants to each session to ensure that they had space for free-flowing conversation.

Most participants arrived at our sessions by microtransit. In each session, participants completed a short survey, followed by 60 minutes of discussion. The participants indicated that they use microtransit for commuting to work (68%), going to doctor's appointments (87%), and running daily errands (74%).

Many participants expressed flexibility in their travel schedules such that they were willing to identify vulnerable others whose rides should be prioritized over their own. Table 1 provides some comments made by focus group participants showing attitudes of prosociality and flexibility, as well as constraints that would be limiting factors for them.

4 SOLUTION CONCEPT AND ILLUSTRATION VIA A MOBILE APP

Figure 2 shows the proposed operation of the entire system, which we dub *Cooperative Adaptive Ride Sharing or CARS.*

This work focuses on the shaded region, developing the CARS agent to understand users (riders) and produce effective and persuasive suggestions for a rider, given the current conditions of the environment, fellow riders, and the agent's knowledge of rider preferences.

We have built a prototype mobile app for microtransit to demonstrate our idea. We use ArcGIS, a collection of online geographic system software [7] to perform the geospatial computations required to calculate candidate alternative locations. We consider multiple riders who request trips on the app. Riders are clustered together based on the similarity of their routes, and an optimal route for sharing rides is computed. We then encourage riders to walk to a pickup point to avoid excessive detours. Riders may have a disability, in which case the algorithm will not suggest any alternative pickup point. We note that all currently operating microtransit

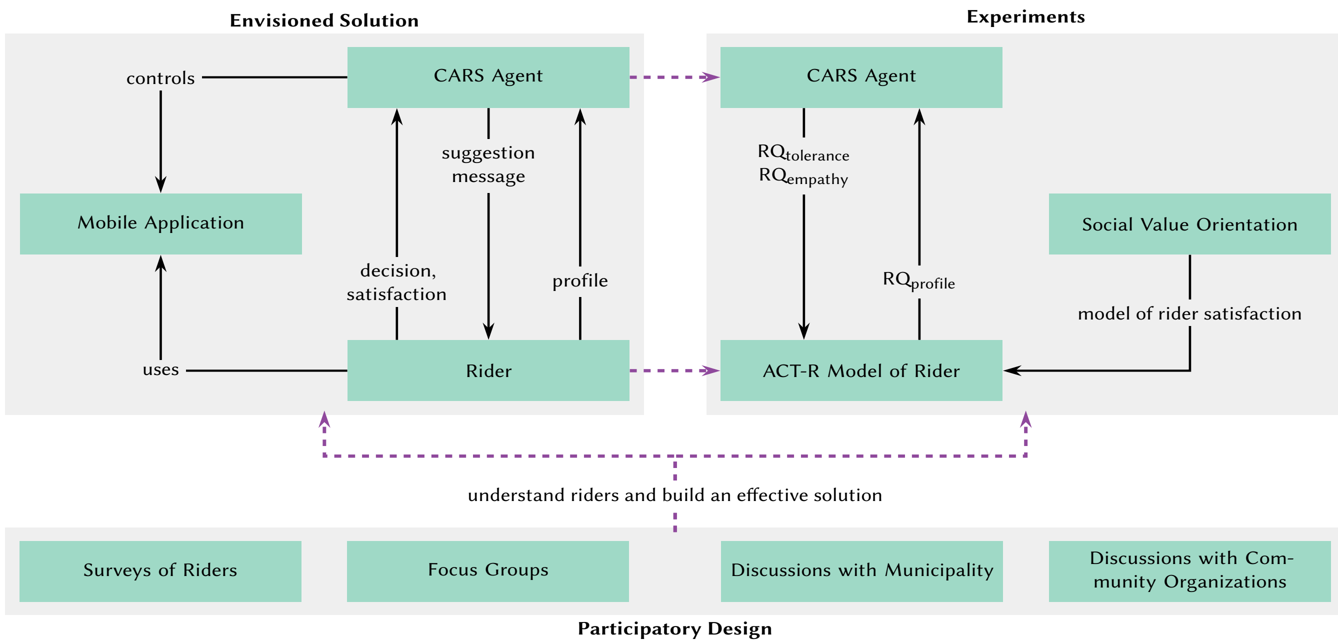


Figure 1: An illustration of our envisioned solution along with the research questions we address.

Table 1: Attitudes of prosociality, constraints, and preferences in microtransit expressed by focus group participants (riders).

Verbatim Comment	Attitude
"Time is on my side, man, I got all the time"	Prosociality, Flexibility: time
"If a person in a wheelchair has a doctor's appointment, then that's a priority"	Prosociality, Other-interest, Empathy: riders in a wheelchair
"Someone who's in our family shelter and has two little kids and a stroller and RIDE doesn't provide car seats, so they have to carry the car seats with them. That's a little hard to walk to any distance"	Prosociality, Empathy: riders with little kids
"Kindness does not cost anything"	Prosociality
"I wouldn't really care. I mean, if someone needs help, I'll try my best to help them out. It's not really a matter of getting something in return"	Prosociality
"We don't need no incentives, just to help somebody out"	Prosociality
"I'm low vision"	Constraint: vision
"I used to have to walk for four, sometimes five blocks, and for somebody who has a bad leg, that's a lot"	Constraint: walking
"I'm sitting here waiting. I'm saying I have disability where I can't stand for a long time. I was at Chiefs [store] and it was raining that day"	Constraint: standing, Preference: avoid rain
"My doctor does want me to walk at times, but not too much"	Constraint: walking
"The doctor, at times, has told me that they want me to get some exercise, so that's about a mile for me to walk per day"	Persuasive factor: health
"Ain't a fan of the rain"	Preference: avoid rain
"Because it's cold"	Preference: avoid cold
"[Young daughter] has to walk from the corner to the house, and it's dark"	Preference: avoid dark
"Going to be stranded" [if wheelchair loses charge]	Constraint: wheelchair
"Work or got doctor's appointments, stuff like that where I need to be on time"	Constraint: urgency

services allow for trip scheduling through a mobile app. However, none of the existing microtransit services enable or encourage users to show flexibility about trip pickup time or location. This app demonstration uniquely incorporates prosocial interventions in a microtransit scheduling platform.

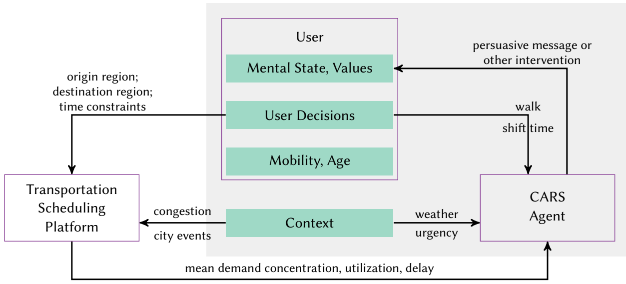
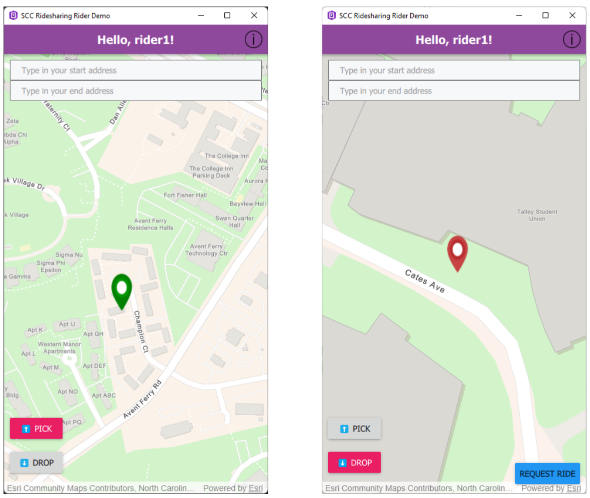


Figure 2: Proposed system operation, highlighting our focus.



(a) Pickup **(b) Dropoff**

Figure 3: Rider pickup and dropoff locations.

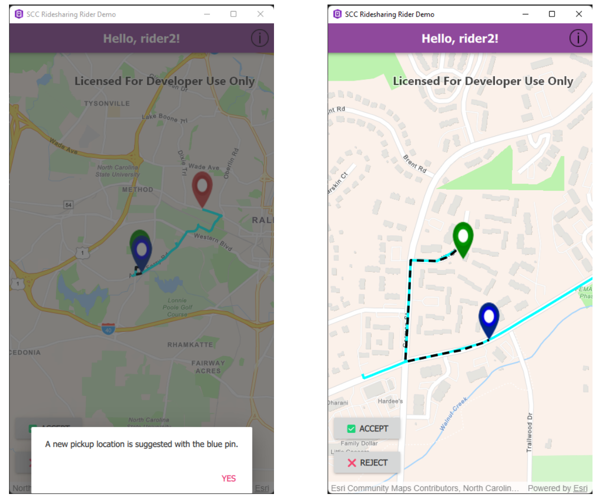
4.1 Rider Pickup and Dropoff Locations

Riders can choose their pickup and dropoff points on the home screen after logging in, as shown in Figure 3. Locations can also be adjusted by moving the green or red pin, respectively. After choosing the pickup and dropoff locations, riders can request a ride by pressing the REQUEST RIDE button on the bottom right of the screen.

4.2 Suggesting an Alternative Location

We calculate the base route by considering the two farthest points in the cluster of pickup and dropoff locations and computing the route between them, considering ordered pairs of [pickup, dropoff] locations. For each rider, we compute the alternative location as the closest point on the base route from their requested location. Riders without disabilities will be suggested this alternative location, which they can accept or reject. The alternative pickup location is shown by a blue pin in Figure 4.

The rider acknowledges this alternative pickup location before proceeding. The walking path between the original pickup point (the green pin) and the alternative pickup point is shown in Figure 4



(a) Alternative Point Calculated **(b) Alternative Point Details Displayed**

Figure 4: Presenting a suggestion to a rider.

(b) as a black dotted line. Riders can choose to accept or reject the alternative location.

If the rider accepts the suggestion, the route is recalculated with the alternative pickup point, as shown in Figure 5 (a). The blue path depicts the final route to be taken by the driver. In future work, we imagine that a rider who accepts a suggestion would accrue KARMA POINTS. Those points could be used to gamify the app: to prioritize riders for timing and convenience (e.g., door-side pickup in times of need).

In case the suggestion is not accepted by the rider, the original path is used. As shown in Figure 5 (b), the route moves into the side road to pick up the rider. This would also happen in the case the rider has a disability, as in that case, the rider is picked up at their requested location.

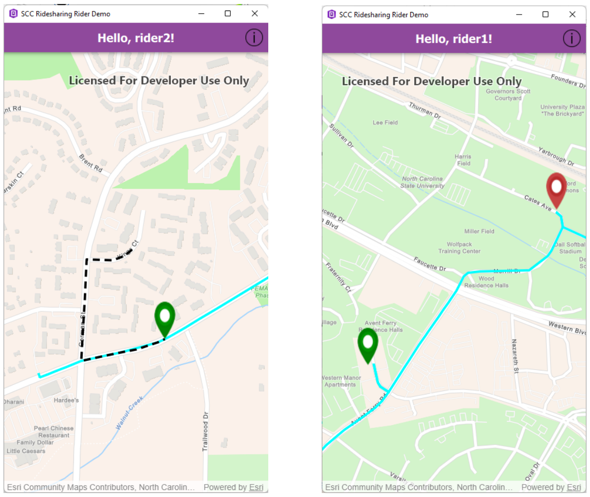
5 CARS AGENT

Rider preferences play a large role in the suggestions they accept. We aim to learn these preferences in two dimensions: a rider's spatial (walking) tolerance under different environmental conditions (contexts), and the persuasive strategies that they respond to. We describe a reinforcement learning approach to learn rider preferences in these two dimensions. In this study, we consider a simplified environment with two features:

- Weather: sunny or rainy
- Time of day: morning, afternoon or evening

5.1 Spatial Adjustment Learning

We use model-free reinforcement learning (Proximal Policy Optimization) to learn optimal spatial suggestions for riders. We experiment with two models (with different reward functions) trained through interactions with the rider, and a customized model trained only on rider profile data.



(a) Accept suggestion **(b) Ignore suggestion**

Figure 5: Rider responding to a suggestion.

We presume that riders have certain spatial tolerances for each of the six environmental combinations introduced in Section 5. We experiment with two reward functions, one that considers only the magnitude of accepted spatial adjustment by the rider, and one that considers rider satisfaction as well. Our aim is to learn how far a rider would be willing to walk under different environmental conditions (contexts). In the following equations, c, s, and r are the context, spatial adjustment suggested, and rider satisfaction with the spatial adjustment s respectively.

5.1.1 Customized Model (PROFILE). This model is trained on rider profile data. Riders provide their ordered preferences in the form of a *feature trace* [6]. A possible weather preference is *sunny* > *rainy*, and a possible time of day preference is *morning, afternoon* > *evening*. We also consider other data provided by the rider (their age and gender). Our aim is to start out with some understanding of the rider with minimal input from them.

An estimation of the rider's spatial tolerance can be calculated as a function of their profile data, as shown in Equation 1.

$$\text{estimated-spatial-tolerance}[c] = F(age, gender, feature\text{-}trace[c]) \tag{1}$$

The customized model is trained to learn these estimated spatial tolerances, as shown in Equation 2.

$$\text{reward}[s, c] = \begin{cases} s, & \text{if } s \leq \text{estimated-spatial-tolerance}[c] \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

5.1.2 Spatial Model (SPATIAL). This model is trained through interactions with the rider and rewarded based on accepted spatial adjustment, as shown in Equation 3.

$$\text{reward}[s, c] = \begin{cases} s, & \text{if } s \text{ is accepted by the rider} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

5.1.3 Rider Satisfaction Model (SAT+SPATIAL). This model is trained through interactions with the rider and rewarded based on accepted spatial adjustment and rider satisfaction, as shown in Equation 4.

$$\text{reward}[s, c] = \begin{cases} s \times r, & \text{if } s \text{ is accepted by the rider} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

5.2 Persuasive Strategy Learning

The second preference we wish to learn is *who* and *what* riders are willing to adjust for. In this study, we consider three persuasive components for the implementation of a persuasive strategy: empathy for those in need, environmental benefit, and health benefit to oneself. We consider six categories of people that could be presumed to be in need of support: *babies*, *children*, *senior citizens*, *ill people*, *people with disabilities*, and *neurodiverse individuals*. People may also empathize with *environmental benefit* and *health benefit* (due to saving fuel costs and walking, respectively). We refer to these categories of people and factors as *value phrases*. We assume that each rider may be empathetic toward a subset of these value phrases.

In our study, we simulate current conditions of the environment during each ride, i.e., the fellow riders and general factors. If the rider rejects the spatial adjustment, we attempt to persuade them further by suggesting who or what they would be helping by making this compromise. We want to learn what is most persuasive to the rider to maximize the chance of them accepting our suggestion. We use a multiarmed bandit algorithm to learn the rider's empathetic preferences based on their response to this second prompt.

6 COGNITIVE MODELING

To build a realistic model of rider decision-making to represent riders in our experiments, we use a cognitive architecture, i.e., a psychological model of human cognition upon which specific tasks can be defined [3]. We adopt the cognitive architecture ACT-R (Adaptive Control of Thought—Rational) [2] because it has an easy-to-use Python library [8]. Accordingly, the riders in our experiments that the CARS agent interacts with are defined using ACT-R. Currently, we do not model drivers in our experiments.

6.1 Overview of ACT-R

In ACT-R [2], the human mind is modeled as a set of three components or cognitive modules, which work together to process information and produce behavior. These three components are the *declarative memory*, which stores facts; the *procedural memory*, which defines actions; and the *control system*, which coordinates the interaction between the declarative and procedural memory and generates behavior according to environmental state.

Decision-making in ACT-R works by way of utilities estimated for different productions. Procedural memory defines the available productions and their prerequisites and consequent states. Given the world state (represented by certain chunks in the goal buffer), we define productions (actions) that can be carried out.

After a production is fired, ACT-R carries out the transitions until the final state, and the utilities of the productions leading up to that state are updated based on the final reward. Over time, the productions which lead to a reward have a higher utility, and hence the ACT-R rider 'prefers' those productions given a choice of multiple productions for a certain set of chunks in the goal buffer.

6.2 Social Value Orientation

Social value orientation (SVO) is a concept in social psychology that states that different individuals have different preferences regarding the allocation of resources between themselves and others [11]. Social value orientation affects behavior in social dilemmas [4]. For the task of adjusting for the sake of others with respect to microtransit rides, social value orientation hence plays a large role, so we use SVO as a part of the internal rewards in the ACT-R riders.

A person's social value orientation is reflected in the internal value they acquire from performing actions. Correspondingly, the internal satisfaction a rider gets from an action (accepting or rejecting an adjustment that benefits someone else) changes based on their social value orientation.

We adopt a simplified version of the classic SVO framework [16], modeling an ACT-R rider's SVO using two parameters.

Other-interest is the degree to which the rider values the outcome of others relative to their own. It ranges from 0 (no interest in the outcome of others) to 1 (only interested in the outcome of others). The complement of other-interest is *self-interest* (i.e., self-interest = 1 − other-interest).

Prosociality is the degree to which the rider values the sum of outcomes for themselves and others. It ranges from 0 (completely competitive) to 1 (completely prosocial). The complement of prosociality is *competitiveness* (i.e., competitiveness = 1 − prosociality).

6.3 Rider Internal Reward Definition

For a realistic utility update of riders' actions, we must define rewards that accurately reflect the internal value a rider gets from a certain action, i.e., the satisfaction they acquire. We model this process as goal-directed choice or value-based decision-making [18]. Riders make decisions based on a comparison of utilities, reflecting the satisfaction they are likely to receive from their choices.

In this study, we assume that riders' internal reward achieved by accepting (different amounts of adjustment) suggestions is inversely proportional to the amount of adjustment. The internal reward attained by rejecting suggestions is directly proportional to the amount of adjustment: riders are likely to feel worse about rejecting interventions that inconvenience them less. In addition, for an acceptance, internal reward is directly proportional to prosociality and other-interest, while for a rejection, internal reward is inversely proportional to prosociality and other-interest.

7 EXPERIMENTS AND RESULTS

In this study, we model riders to have certain spatial tolerances for each of the possible contexts ([weather, time of day] combinations) mentioned in Section 5, as well as certain value phrases they are empathetic toward, as mentioned in Section 5.2. We aim to learn both these preferences. Our CARS agent combines this knowledge to persuade riders to accept spatial adjustments and behave prosocially. To assess the performance of our approach in rider modeling, we run experiments with five diverse riders.

Table 2: Rider personas for this study.

SVO	Gender	Age	Other-interest	Prosociality
Competitive	Male	Early 30s	0.01	0.01
Individualistic	Female	16	0.01	0.50
Moderate	Female	Mid 40s	0.45	0.50
Prosocial	Male	Early 50s	0.50	0.99
Altruistic	Female	Early 70s	0.99	0.99

7.1 Rider Personas

We consider four archetypal personalities (competitive, individualistic, prosocial, and altruistic) according to SVO research as well as one moderate persona to offset the extremities. Competitive people seek to maximize the difference between their outcomes and those of others. Individualistic people are concerned only with maximizing their own outcomes. Prosocial people prefer mutually beneficial outcomes, while altruistic people are interested only in the outcomes of others. Our moderate persona prioritizes their outcomes slightly more than others', and is neither prosocial nor competitive. A summary of our rider personas is shown in Table 2. We show sample rider profile data and internal preferences (used for the prosocial rider in our experiments) in Tables 3 through 5.

7.2 Hypotheses

We refine our research questions into these evaluable claims.

$H_{tolerance}$ *Spatial tolerances for riders can be learned over time.* This knowledge can be used to suggest optimal spatial adjustments that would benefit the system without causing too much inconvenience to riders.

$H_{empathy}$ *Persuasive strategies to promote prosociality can be learned over time.* This can be used to persuade riders to behave prosocially even if the spatial adjustment causes them some inconvenience.

$H_{profile}$ *Knowledge about a rider's basic profile will provide a better (nonnaive) starting point to interact with the rider*, i.e., we can use basic rider profile data to initialize a nonnaive (customized) model to start with.

7.3 Evaluation Metrics

We evaluate the performance of the CARS agent as a combination of the performance of the spatial tolerance learning and the persuasive strategy learning aspects of it.

7.3.1 CARS Agent–Spatial Tolerance Learning. We use the *average accepted spatial adjustment per episode* and the *average acceptance percentage per episode* as our evaluation metrics. We evaluate the three models mentioned in Section 5.1.

Table 3: Sample rider profile: Prosocial rider

Parameter	Value
Gender	Male
Age	Early 50s
Weather feature trace	Sunny > Rainy
Time of day feature trace	[Morning = Evening] > Afternoon

Table 4: Sample spatial tolerances: Prosocial rider

Weather	Time of Day	Spatial Tolerance (in meters)
Sunny	Morning	475
Sunny	Afternoon	200
Sunny	Evening	400
Rainy	Morning	250
Rainy	Afternoon	200
Rainy	Evening	20

Table 5: Sample empathy (persuasive value phrase) distribution: Prosocial rider

Value Phrase	Persuasive Percentage
Environmental benefit	51%
People with disabilities	49%

7.3.2 CARS Agent–Persuasive Strategy Learning. We evaluate the performance of our agent by calculating the similarity between the actual and predicted value phrase distributions for the rider.

7.4 Results

We show results for spatial adjustment learning for all the riders in Figures 6, 7, 8, and 9. We show the difference between actual and predicted persuasive value phrase distributions in Table 6, and Figure 10 shows the actual versus top three predicted value phrase percentages for the prosocial rider.

7.4.1 Evaluating Hypothesis $H_{tolerance}$. Figures 6 and 7 show spatial adjustment results in terms of the average accepted adjustment and acceptance percentage for each rider for the three models and the untrained baseline, measured over 500 episodes of 1,000 time steps each. There is a slight improvement in PROFILE compared to the untrained model, indicating the benefit of customization with the rider profile. SPATIAL, which has been trained for 150,000 time steps with the rider, performs only slightly better than PROFILE. However, SAT+SPATIAL, which has been trained for the same number of steps with the rider, performs better than the other models in terms of both average accepted adjustment and average acceptance percentage. This indicates that considering rider satisfaction is beneficial for the system as well, as riders can be persuaded to make larger adjustments if we consider their satisfaction while accepting adjustments. The improvement in SAT+SPATIAL is much more pronounced for the Competitive, Individualistic, and Moderate riders, indicating that even riders who are proself can behave prosocially if we understand them better. Most people are not completely prosocial (as are the Prosocial and Altruistic personas, who, being completely prosocial, accept almost all suggestions even if they are highly inconvenienced by them), and so are likely to be much more satisfied (and willing to adjust) if we take their satisfaction into account.

Figures 8 and 9 show the Cohen's d effect sizes and corresponding 95% confidence intervals of SAT+SPATIAL compared to SPATIAL, PROFILE, and the untrained baseline.

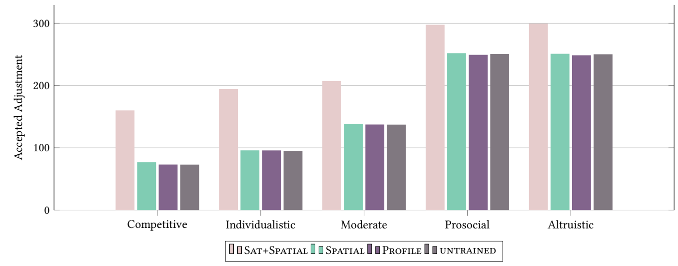


Figure 6: Average Accepted Adjustment for the three models and the baseline (untrained).

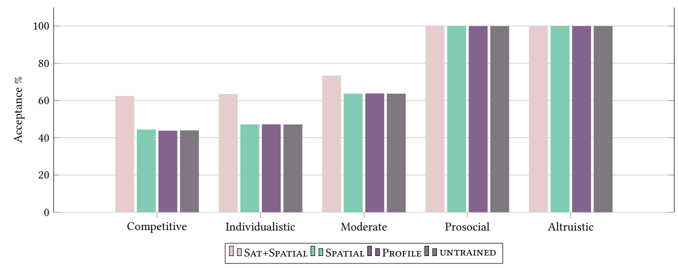


Figure 7: Average Acceptance Percentage for the three models and the baseline (untrained).

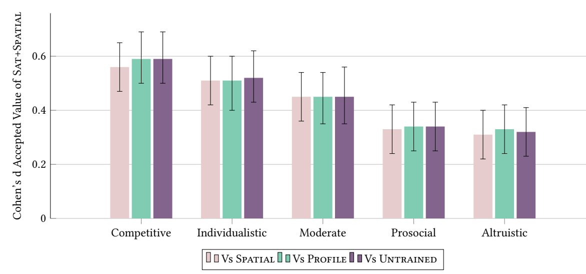


Figure 8: Effect sizes measured as Cohen's d scores and corresponding 95% confidence intervals of the improvement of SAT+SPATIAL over SPATIAL, PROFILE and an untrained model for accepted adjustment value.

7.4.2 Evaluating Hypothesis $H_{empathy}$. Figure 10 provides a summary of the persuasive value phrase distribution learned by our bandit model for the prosocial rider, and Table 6 shows the Hellinger distance between the actual and predicted persuasive value phrase distributions for each of the experiments. The Hellinger distance

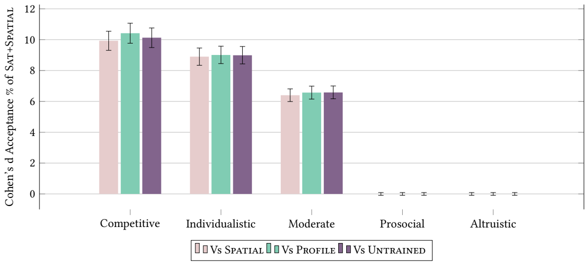


Figure 9: Effect sizes measured as Cohen's d scores and corresponding 95% confidence intervals of the improvement of Sat+Spatial over Spatial, Profile and an untrained model for the percentage of accepted suggestions.

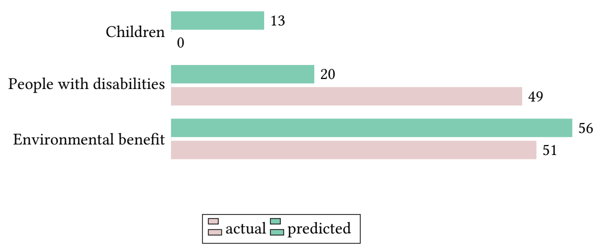


Figure 10: Prosocial rider: actual vs top three predicted persuasive value phrase percentages.

Table 6: Hellinger distance between actual and predicted persuasive value phrase distributions.

Experiment	Hellinger Distance
Prosocial	0.391
Altruistic	0.338
Moderate	0.604
Competitive	0.316
Individualistic	0.261

quantifies the difference between two probability distributions and is defined for two discrete probability distributions $P = (p_1, \ldots, p_k)$ and $Q = (q_1, \ldots, q_k)$ as

$$\text{H(P,Q)} = \frac{1}{\sqrt{2}} \times \sqrt{\sum_{i=1}^{k} (\sqrt{p_i} - \sqrt{q_i})^2}. \tag{5}$$

7.4.3 Evaluating Hypothesis $H_{profile}$. With the current rider profile data, there is a slight improvement in Profile compared to the untrained baseline, as shown in Figures 6 and 7.

7.4.4 Summarized Results. We now summarize how we answer the research questions introduced in Section 1.

RQ$_{\textbf{tolerance}}$ *Can we learn riders' spatial tolerances to suggest optimal spatial adjustments?* We answer this question positively based on Figures 6, 7, 8, and 9 in Section 7.4.1.

RQ$_{\textbf{empathy}}$ *Can we learn riders' empathetic tendencies to persuade them to adjust?* We answer this question positively based on Table 6 in Section 7.4.2 and Figure 10 in Section 7.4.2.

RQ$_{\textbf{profile}}$ *Could considering rider profile data lead to a better (non-naive) starting point?* We answer this question positively based on Figures 6 and 7 in Section 7.4.1.

8 DISCUSSION

We present a conception of a prosocial approach to microtransit to create a more equitable and sustainable ecosystem. We demonstrate the working of our idea with a prototype app, as well as experiments using a cognitive architecture as a surrogate for a human rider, to show that rider preferences (both acceptable spatial adjustments and persuasive strategies) can be learned with reinforcement learning and used to persuade them to help others. We show that basic rider profile data is enough to customize a model with some initial knowledge.

We find that the methods are effective for riders with varying levels of prosociality. As we found through the focus groups, many riders are prosocial at the outset, and considering rider satisfaction can help persuade even the less prosocial riders to behave prosocially.

We show that with ACT-R, we are able to model diverse riders, whose varied responses to suggestions indicate the differing internal satisfaction they derive from their actions. Our results show that ACT-R can be used to model human decision-making in simulations where human input is required and relevant or sufficient data is not available, accounting for people with different behaviors and motivations.

Our results suggest that if we combine this work with an optimizer that optimizes a system-level or higher priority metric, we could calculate adjustments that are of the least inconvenience to riders while increasing the prosociality of the entire system.

8.1 Challenges

Using AI-based interventions to change user preferences or behavior, even for a good societal objective, is potentially ethically risky. Key challenges include a deeper understanding of consent [23] and privacy requirements [17] so that an AI agent does not violate a user's autonomy.

A more general challenge is that of achieving trust. A decision about trust brings forth judgments of an agent's ability, benevolence, and integrity [14]. The same constructs form an effective basis for assessing the trustworthiness of AI agents [21]. Besides the inherent benefits of ensuring that our STSs promote trust and that our agents are trustworthy, another motivation for trust is practical: Once a user loses trust in the system, they may elect not to participate or participate only to the extent necessary, e.g., by disregarding any attempted persuasion and thereby forgoing the prosocial outcomes we desire.

ACKNOWLEDGMENTS

We thank Aditya Khadse for his effort in implementing our prototype. We thank the anonymous reviewers for their helpful comments. We acknowledge support from the US National Science Foundation (grants SCC-2125447 and SCC-2325720).

REFERENCES

[1] Evangelia Anagnostopoulou, Efthimios Bothos, Babis Magoutas, Johann Schrammel, and Gregoris Mentzas. 2018. Persuasive Technologies for Sustainable Mobility: State of the Art and Emerging Trends. *Sustainability* 10, 7 (2018), 2128. https://doi.org/10.3390/su10072128

[2] John Robert Anderson. 1996. ACT: A simple theory of complex cognition. *American Psychologist* 51, 4 (1996), 355–365. https://doi.org/10.1037/0003-066X.51.4.355

[3] John R. Anderson and Christian Lebiere. 2003. The Newell Test for a Theory of Cognition. *Behavioral and Brain Sciences* 26, 5 (2003), 587–601. https://doi.org/10.1017/S0140525X0300013X

[4] Daniel Balliet, Craig Parks, and Jeff Joireman. 2009. Social Value Orientation and Cooperation in Social Dilemmas: A Meta-Analysis. *Group Processes & Intergroup Relations* 12, 4 (2009), 533–547. https://doi.org/10.1177/1368430209105040

[5] Eleni Bardaka, Leila Hajibabai, and Munindar P. Singh. 2020. Reimagining Ride Sharing: Efficient, Equitable, Sustainable Public Microtransit. *IEEE Internet Computing (IC)* 24, 5 (Sept. 2020), 38–44. https://doi.org/10.1109/MIC.2020.3018038

[6] Andreea Bobu, Marius Wiggert, Claire J. Tomlin, and Anca D. Dragan. 2022. Inducing Structure in Reward Learning by Learning Features. *International Journal of Robotics Research* 41, 5 (April 2022), 497–518. https://doi.org/10.1177/02783649221078031

[7] Bob Booth, Andy Mitchell, et al. 2001. *Getting Started with ArcGIS*. ESRI New York.

[8] Adrian Brasoveanu and Jakub Dotlačil. 2020. The ACT-R Cognitive Architecture and Its pyactr Implementation. In *Computational Cognitive Modeling and Linguistic Theory*. Springer International Publishing, Chapter 2, 7–37. https://doi.org/10.1007/978-3-030-31846-8_2

[9] Lucia Cipolina-Kun, Sebastian Stein, Vahid Yazdanpanah, and Enrico Gerding. 2022. Coalition Formation in Ridesharing with Walking Options. In *ICLR 2022 Workshop on Gamification and Multiagent Solutions*. https://openreview.net/forum?id=BcgYj9kaeq Accessed 2023-06-28.

[10] Christine Kormos, Reuven Sussman, and Bracha Rosenberg. 2021. How cities can apply behavioral science to promote public transportation use. *Behavioral Science & Policy* 7, 1 (2021), 95–115. https://doi.org/10.1353/bsp.2021.0004

[11] Paul A. M. Van Lange. 1999. The Pursuit of Joint Outcomes and Equality in Outcomes: An Integrative Model of Social Value Orientation. *Journal of Personality and Social Psychology* 77, 2 (1999), 337. https://doi.org/10.1037/0022-3514.77.2.337

[12] Catherine Marshall and Gretchen B. Rossman. 2021. *Designing Qualitative Research*. SAGE Publications, Inc. https://books.google.com/books?id=wzwiEAAAQBAJ Accessed 2023-06-28.

[13] Joseph A. Maxwell. 2012. *Qualitative Research Design: An Interactive Approach*. SAGE Publications, Inc. https://books.google.com/books?id=xAHCOmtAZd0C Accessed 2023-06-28.

[14] Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model of Organizational Trust. *The Academy of Management Review* 20, 3 (July 1995), 709–734. https://doi.org/10.5465/amr.1995.9508080335

[15] David L. Morgan. 2019. *Basic and Advanced Focus Groups*. SAGE Publications, Inc. https://doi.org/10.4135/9781071814307

[16] Ryan O. Murphy and Kurt A. Ackermann. 2014. Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences. *Personality and Social Psychology Review* 18, 1 (2014), 13–41. https://doi.org/10.1177/1088868313501745

[17] Gideon Ogunniye and Nadin Kokciyan. 2023. A Survey on Understanding and Representing Privacy Requirements in the Internet-of-Things. *Journal of Artificial Intelligence Research* 76 (2023), 163–192. https://doi.org/10.1613/jair.1.14000

[18] Antonio Rangel. 2009. The Computation and Comparison of Value in Goal-directed Choice. In *Neuroeconomics*, Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, and Russell A. Poldrack (Eds.). Academic Press, London, Chapter 28, 425–440. https://doi.org/10.1016/B978-0-12-374176-9.00028-2

[19] Sören Schleibaum and Jörg P. Müller. 2020. Human-Centric Ridesharing on Large Scale by Explaining AI-Generated Assignments. In *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*. Association for Computing Machinery, 222—-225. https://doi.org/10.1145/3411170.3411274

[20] Susan Shaheen, Adam Cohen, Nelson Chan, and Apaar Bansal. 2020. Sharing strategies: Carsharing, shared micromobility (bikesharing and scooter sharing), transportation network companies, microtransit, and other innovative mobility modes. *Transportation Land Use and Environmental Planning, Berlin, Germany: Elsevier* 1 (2020), 237–262. https://doi.org/10.1016/B978-0-12-815167-9.00013-X

[21] Amika M. Singh and Munindar P. Singh. 2023. Wasabi: A Conceptual Model for Trustworthy Artificial Intelligence. *IEEE Computer* 56, 2 (Feb. 2023), 20–28. https://doi.org/10.1109/MC.2022.3212022

[22] Munindar P. Singh. 2014. Norms as a Basis for Governing Sociotechnical Systems. *ACM Transactions on Intelligent Systems and Technology* 5, 1, Article 21 (Jan. 2014), 23 pages. https://doi.org/10.1145/2542182.2542203

[23] Munindar P. Singh. 2022. Consent as a Foundation for Responsible Autonomy. *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)* 36, 11 (Feb. 2022), 12301–12306. https://doi.org/10.1609/aaai.v36i11.21494 Blue Sky Track.

[24] Sebastian Stein and Vahid Yazdanpanah. 2023. Citizen-Centric Multiagent Systems. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2023, London, United Kingdom, 29 May 2023 - 2 June 2023*, Noa Agmon, Bo An, Alessandro Ricci, and William Yeoh (Eds.). Association for Computing Machinery, 1802–1807. https://doi.org/10.5555/3545946.3598843

[25] Divya Sundaresan, Akhira Watson, Eleni Bardaka, Crystal Chen Lee, Christopher B. Mayhorn, and Munindar P. Singh. 2025. Prosociality in Microtransit. *Journal of Artificial Intelligence Research (JAIR)* 82 (Jan. 2025), 77–110. https://doi.org/10.1613/jair.1.16777

[26] Jamal Yousaf, Juanzi Li, Lu Chen, Jie Tang, and Xiaowen Dai. 2014. Generalized multipath planning model for ride-sharing systems. *Frontiers of Computer Science* 8 (2014), 100–118. https://doi.org/10.1007/s11704-013-3021-6