Safety21
INNOVATING FOR SAFETY

US DOT National
University Transportation Center for Safety

Carnegie Mellon University

Project Title

Planning and Policy for Safer Roads with Autonomous Vehicles: Moral Decision Making Behavior in Dilemma-inducing Situations

Investigators

Fatemeh Nazari (https://orcid.org/0000-0003-1587-1848)
Mohamadhossein Noruzoliaee (https://orcid.org/0000-0003-3860-8911)

Final Report – August 2025

DISCLAIMER

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
549</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Planning and Policy for Safer Roads with Autonomous Vehicles: Moral Decision Making Behavior in Dilemma-inducing Situations</td><td colspan="2">5. Report Date
August 19, 2025</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Fatemeh Nazari, Ph.D. https://orcid.org/0000-0003-1587-1848
Mohamadhossein Noruzoliaee, Ph.D. https://orcid.org/0000-0003-3860-8911</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
The University of Texas Rio Grande Valley (UTRGV)
1201 W University Dr
Edinburg, Texas 78539</td><td colspan="2">10. Work Unit No.</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
Federal Grant # 69A3552344811 / 69A3552348316</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Safety21 University Transportation Center
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213</td><td colspan="2">13. Type of Report and Period Covered
Final Report (July 1, 2024 – June 30, 2025)</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
USDOT</td></tr>
<tr><td colspan="4">15. Supplementary Notes
Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.</td></tr>
<tr><td colspan="4">16. Abstract
This study develops a Dynamic Bayesian Network (DBN) framework to examine how public confidence in autonomous vehicle (AV) safety and willingness-to-ride respond to policy interventions in crash-imminent pedestrian–passenger prioritization scenarios. Using stated preferences survey data from San Francisco (SF) and San Antonio (SA), the model integrates baseline attitudes, socio-demographic factors, and policy conditions to simulate both intra-slice and inter-slice dependencies. Results from empirical-mix simulations indicate that SF respondents, despite having higher baseline confidence and willingness-to-ride, exhibit greater sensitivity to policies prioritizing pedestrians, with significant declines across all scenarios. By contrast, SA respondents show comparatively stable and modestly positive shifts, particularly when policies favor passengers. Stratified analyses reveal heterogeneity so that policies such as pedestrian prioritization amplify existing differences across baseline confidence groups, while others, like prioritization by child presence or group size, promote convergence toward midscale attitudes. The findings underscore the value of DBNs in capturing causal and temporal dynamics in AV acceptance and highlight important city-level and attitudinal differences in policy responsiveness.</td></tr>
<tr><td colspan="2">17. Key Words
Autonomous vehicle; Safety Perceptions; Dynamic Bayesian Networks; Travel Behavior; Policy Simulation</td><td colspan="2">18. Distribution Statement
No restrictions.</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
34</td><td>22. Price</td></tr>
</table>

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

1. Introduction

The rapid advancements in vehicle autonomy and artificial intelligence (AI) have propelled autonomous vehicles (AVs) from experimental prototypes to technologies undergoing extensive road testing (Traffic Lab, 2024) and, more recently, limited commercial deployment (Waymo, 2024). While these advances signal the potential for a transformative shift in personal and shared mobility, a critical challenge to integration of AV into the real-world transportation systems as well as our personal and professional lives is establishing regulatory and policy frameworks that define AV behavior in complex situations. In particular, such principles would determine how an AV makes decisions in dilemma-inducing situations, for instance, whether an AV should prioritize the safety of the onboard passengers over that of pedestrians in unavoidable collision circumstances.

The present study approaches the decision-making challenge not from a purely philosophical standpoint, which seeks to prescribe normative rules for AV behavior (Geisslinger et al., 2021), but rather from the perspective of public expectations and perceptions which is at the intersection of economics, psychology, and cognitive science. These disciplines offer a lens into the processes underlying human decision-making, particularly when choices carry consequences for others as well. Building on this understanding, the present research seeks to examine how the public envisions AV decision-making by investigating the cognitive and attitudinal factors that shape these expectations. The findings aim to inform AV policy and system design, ensuring that decision-making frameworks align with both technical feasibility and widely held public preferences.

Research on AV decision-making has increasingly moved beyond purely technical considerations to explore how public expectations, socio-political acceptance, and behavioral realities should inform AV policy and design. This literature reflects a shift from normative, philosophy-driven prescriptions toward empirical studies that capture how people actually reason about safety trade-offs and risk in both hypothetical and real-world contexts. In reviewing the most relevant works, we focus on empirical and experimental findings as well as review-based contributions that illuminate how individuals navigate safety dilemmas, and how these preferences might be embedded into AV decision frameworks.

From a human-centered systems standpoint, Xing et al. (2021) provide a comprehensive review of collaborative autonomous driving systems, emphasizing the pivotal role of cognitive factors (such as situational awareness, shared control strategies, and trust calibration) in enabling effective human–AV interaction. They show that systems incorporating real-time driver state monitoring, which captures indicators like attentional focus, readiness, and cognitive load, substantially enhance public acceptance of AVs. By analyzing two principal interaction modes, shared control (where human and autonomy jointly manage driving) and takeover control (where responsibility transfers entirely between agent and driver), the authors highlight the need for interfaces that support smooth transitions and maintain operator understanding.

At the policy and industry level, Martinho et al. (2021) challenge the prevailing academic emphasis on extreme case dilemmas by examining real-world priorities within the AV industry. Through a systematic review of corporate reports and technical publications from AV developers in California, the authors reveal that industry attention is heavily concentrated on safety, cybersecurity, liability mitigation, and regulatory compliance, rather than philosophical thought experiments such as the trolley problem. While the academic literature is dominated by hypothetical scenarios, AV companies appear to address those dilemmas only indirectly or implicitly, instead opting for practical design strategies that minimize risk, expedite incident investigations, and favor robust crash avoidance systems.

Complementing these perspectives, Bergmann et al. (2018) explore the essential role of socio-political acceptance in shaping AV decision-making policies. Through experimental philosophy, the authors present real-world driving dilemmas, revealing that participants demonstrate a willingness to sacrifice themselves to save others and consider the age of potential victims when making split-second decisions, even when such actions may contravene traffic norms. These findings underscore that public expectations balance self-preservation with broader safety, suggesting that AV algorithms must not only reflect technical feasibility but also align with social intuitions, especially in situations requiring judgment.

Radun et al. (2019) moves beyond theoretical trolley-type dilemmas by examining a real-world scenario often faced by drivers as whether veering off-road to avoid a potentially fatal collision with a lighter oncoming vehicle would risk a rollover that could endanger the driver's safety. This real-life dilemma highlights real-world stakes and complexities that standard thought experiments overlook. Their experimental survey reveals that, in practice, drivers show a clear reluctance to endanger themselves for the sake of others. Crucially, this finding underscores that public decision-making in real-world scenarios is shaped not by abstract rules but by concrete risk assessment and personal safety considerations. Krügel and Uhl (2022) challenge the conventional framing of AV regulations, which often centers on stark binary dilemmas with zero uncertainty. Across three online experiments, they reveal that the public does not simply favor minimizing accident risk at all costs. Instead, respondents prefer a nuanced balancing of multiple harms in real-world decision-making contexts. This indicates that preferences in such situations emerge differently when risky scenarios involve trade-offs rather than clear-cut, high-stakes outcomes.

The empirical study by Sui (2023) provides critical insights into how the public values different algorithms for AV decision-making. Through an online survey of 460 participants in China, the study assesses preferences across four algorithmic approaches, utilitarianism, rawlsianism, egoism (favoring self-interest), and a hybrid model, using a combination of trolley-like dilemma scenarios and Likert-scale acceptability ratings. The results challenge simplistic assumptions so that both pure egoism and pure utilitarianism face similarly high rejection rates, as participants disfavor algorithms that either promote complete self-sacrifice or prioritize the greater good at personal expense. Instead, a hybrid strategy, which balances harm minimization with self-responsibility, emerges as the most endorsed, indicating public preference for context-sensitive frameworks over rigid paradigms.

The present study aims to address the above gaps in understanding public attitudes toward AV decision-making by developing a data-driven Dynamic Bayesian Network (DBN) model. Unlike normative approaches that prescribe how AVs should behave, our framework adopts a descriptive perspective, capturing how the public actually envisions AV behavior in realistic, context-dependent scenarios. The model is calibrated using empirical data from two distinct urban populations and simulates scenario-specific attitudes toward confidence in AV safety and willingness-to-ride with AV, accounting for variations in baseline confidence and situational context. The findings offer actionable insights on multiple fronts. First, the results provide policymakers, industry stakeholders, and transportation planners with evidence-based guidance for designing AV policies and systems that align with public expectations. Second, the model advances travel behavior research by integrating cognitive and attitudinal factors, such as trust and acceptance, into AV adoption analysis. Finally, by delivering a scalable, empirically grounded framework, this work contributes to broader AI governance discussions, offering a transparent and adaptable tool for aligning autonomous decision-making systems with user-centered priorities.

The remainder of the report is structured as follows. Section 2 presents the study methodology, including the DBN framework. Section 3 details the survey design and data collection process, followed by Section 4 analyzing the empirical dataset. Section 5 reports the simulation results from calibrating the DBN model under alternative policy scenarios. Finally, Section 6 concludes with a summary of key findings, limitations, and directions for future research.

2. Methodology

Bayesian Networks (BNs) have gained significant attention as a powerful analytical framework for modeling complex interdependencies across domains such as economics, medicine, environmental science, and transportation (Pearl, 2009). A BN is a directed probabilistic graph in which nodes represent variables and edges encode conditional dependencies, offering an intuitive, visual representation of how factors relate to one another. Foundational work by Pearl (1988) established the theoretical basis for BNs, with comprehensive summaries and methodological advancements provided by Friedman et al. (1997), Koski and Noble (2011), and Scutari and Denis (2021). Unlike traditional regression models, BNs do not require pre-specified functional forms, enabling the discovery of key interactions and interdependencies directly from the data. BNs also provide a formal framework for causal reasoning, allowing researchers to simulate how changes in one factor can propagate through the network to influence others, which provides an essential capability for realistic scenario analyses in AV policy research. Additional advantages include the ability to integrate expert knowledge with empirical data and to handle missing data through probabilistic reasoning, thereby reducing bias and preserving the effective sample size (Liu and Motoda, 2012; Madden, 2009; Oh et al., 2022; Jo et al., 2023). Traditional BNs can be extended to DBN by incorporating temporal dependencies, enabling the identification of both contemporaneous and temporal relationships among variables (Dagum et al., 1992). This allows DBNs to capture interactions at a single time point as well as the evolution of these relationships over time.

Building on these concepts, Figure 1 presents the overarching framework for DBN modeling of attitudes towards AV. This framework operationalizes the DBN approach for our study, linking baseline attitudes to post-policy perceptions while integrating socio-demographic and behavioral predictors. The two "time slices" correspond to baseline perceptions before policy implementation ($t = 0$) and post-policy perceptions under a given scenario ($t = 1$). Each node in the network represents either an attitudinal measure — namely, confidence in AV safety ($CONF_t$) and willingness-to-ride (WTR_t) — or a relevant predictor such as socio-demographic characteristics, travel behavior indicators, and the applied policy scenario (P_s). Within each time slice, directed edges represent conditional dependencies among concurrent variables. For example, at baseline ($t = 0$), confidence in AV safety and willingness-to-ride are allowed to influence one another, reflecting the empirical link between perceived safety and readiness to use AV technology. Temporal transitions are modeled through inter-slice dependencies, such that post-policy confidence ($CONF_1$) is conditionally dependent on its baseline value ($CONF_0$), baseline willingness-to-ride (WTR_0), and the policy scenario(P_s). Similarly, post-policy willingness-to-ride (WTR_1) depends on its baseline value (WTR_0), baseline confidence ($CONF_0$), and the scenario. This design captures both autocorrelation (attitude persistence) and cross-effects between confidence and willingness over time.

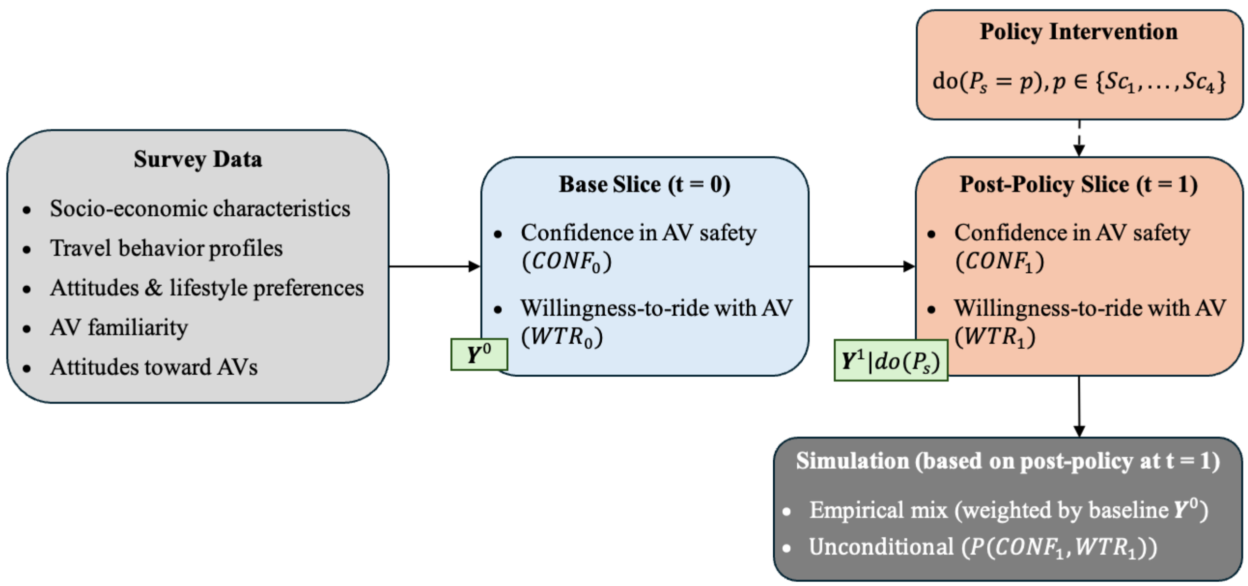


Figure 1. Overarching Framework for Dynamic Bayesian Network-Based Modeling of Confidence in AV Safety and Willingness-to-Ride

Building on the structure outlined in Figure 1, the DBN mathematically formalizes the relationships among baseline and post-policy variables through a joint probability distribution, factored according to the network's conditional independencies. Let $\boldsymbol{Y}^0$ denote all variables at $t = 0$ and $\boldsymbol{Y}^1$ all variables at $t = 1$. The model factorizes as written in Eq. (1), where $\mathrm{Pa}(Y_i^1)$ denotes the set of parents of Y_i^1 in the DBN, including any baseline predecessors. Parameters are estimated using maximum likelihood from the observed baseline survey data. Equation (1) can be expanded into Eqs. (2) and (3), where Y_j^0 and Y_i^1 respectively indicate baseline variable j at $t = 0$ and post-policy variable i at $t = 1$. The first product term in Eq. (2) represents the factorization of the baseline joint probability according to intra-slice dependencies at $t = 0$ and the second product term represents the transition model, factoring post-policy variables given their intra-slice and inter-slice parents. In the present study, the AV attitudinal change component in Eq. (2) is expressed as written in Eq. (3).

$$P(\boldsymbol{Y}^0, \boldsymbol{Y}^1) = P(\boldsymbol{Y}^0) \prod_i P(Y_i^1 | \mathrm{Pa}(Y_i^1)) \tag{1}$$

$$P(\boldsymbol{Y}^0, \boldsymbol{Y}^1) = \underbrace{\left[\prod_j P(Y_j^0 | \mathrm{Pa}(Y_j^0))\right]}_{Baseline\ slice\ (t=0)} \times \underbrace{\left[\prod_i P(Y_i^1 | \mathrm{Pa}(Y_i^1))\right]}_{Post-policy\ slice\ (t=1)} \tag{2}$$

$$\begin{aligned} P(\boldsymbol{Y}^0, \boldsymbol{Y}^1) = {} & P(CONF_0 | WTR_0) \\ & \cdot P(WTR_0 | CONF_0) \times P(CONF_1 | CONF_0, WTR_0, P_s) \\ & \times P(WTR_1 | WTR_0, CONF_0, P_s) \end{aligned} \tag{3}$$

On the right side of the overarching framework in Figure 1, policy interventions are incorporated into the DBN through the causal do-operator, namely, $\mathrm{do}(P_s)$. This operator fixes the policy variable to a specified value while removing the influence of its natural parents. This enables counterfactual simulation of post-policy distributions $P(\boldsymbol{Y}^1 | \mathrm{do}(P_s))$ without requiring actual policy implementation. By isolating the causal impact of the policy scenario, this approach enables the evaluation of how interventions would shift attitudes toward AV safety and adoption, supporting robust, evidence-based policy analysis. Once the DBN is fitted, policy impacts are simulated by propagating probabilities forward from the baseline slice to the post-policy slice under each policy scenario. For each respondent, we compute the conditional probability distribution over confidence or willingness-to-ride levels at $t = 1$, given their baseline characteristics and the imposed policy. These distributions are then aggregated to obtain the expected level $\mathbb{E}[Level]$, the mid-scale probability ($\mathbb{E}[mid\%]$), and the full distribution across the five response levels. The change from baseline is calculated as the difference in expected levels between $t = 1$ and $t = 0$, and statistical significance of scenario effects is assessed using Mann–Whitney (MW) U tests. This process yields scenario-specific outcome distributions for the full sample and, when disaggregated by baseline confidence strata, reveals heterogeneity in policy responsiveness.

3. Survey Design and Data Collection

To collect the sample dataset, we developed a structured survey instrument administered in two distinct U.S. metropolitan regions, namely, San Francisco, California (SF) and San Antonio, Texas (SA). The survey design is informed by prior studies on AV acceptance, decision-making under risk, and stated preferences methods (see reviews in Becker and Axhausen (2017), Gkartzonikas and Gkritza (2019), Harb et al. (2021), and Othman (2021)), ensuring both theoretical grounding and empirical comparability. Questionnaire covered socio-economic characteristics, travel patterns, general attitudes and lifestyle preferences, and AV-related attitudes under baseline and policy-scenario conditions. The attitudinal modules incorporated validated psychometric scales, including items capturing cognitive style dimensions (e.g., emotion- versus reason-driven decision-making) and rule-based orientation (e.g., utilitarian versus deontological tendencies). Scenario-based items were used to elicit respondent judgments in hypothetical AV decision contexts, enabling analysis of how baseline confidence in AV safety and contextual framing influence responses.

Data were collected through an online survey platform in 2025. The survey participants are recruited from blended online market-research panels using incentives such as cash, airline miles, and gift cards. The final dataset comprises 176 valid responses from SF and 159 from SA, allowing for cross-regional comparisons between a technology-forward urban environment and a more traditionally car-oriented metropolitan area. This dual-city sampling strategy was designed to capture geographic and cultural variation in AV-related attitudes, supporting the study's focus on modeling heterogeneity in public expectations of AV behavior using a DBN framework.

4. Data Description

This section presents an overview of the two sample datasets from SF and SA. We first outline the socio-demographic profile of participants, followed by their travel behaviors, attitudinal and lifestyle preferences, and attitudes toward AV technology. Finally, we describe how measures of AV familiarity and attitudes, both in baseline and alternative policy scenarios (i.e., confidence in AV safety and willingness to ride), are operationalized to drive simulations in our DBN model.

4.1. Socio-demographic Characteristics

The sample individuals are characterized by socio-demographic factors, which are statistically distributed as presented in Table 1. Gender distribution is more balanced in SF, whereas SA has a higher proportion of females (61% versus 45%). Educational attainment differs sharply, as SF participants are more likely to hold a bachelor's or advanced degree (68% combined) compared to only 32% in SA, where two-thirds fall into the lower education category. Income patterns follow a similar trend, with high-income households (≥\$150K) far more prevalent in SF (24%) than SA (4%), while low-income households dominate in SA (74% vs. 43%). SF households are also more likely to own alternative fuel vehicles (39% versus 9%). These differences suggest distinct socio-economic contexts, which may influence AV-related attitudes and decision-making preferences across the two cities.

Table 1. Distribution of Socio-economic Characteristics in San Francisco and San Antonio

Variables	Category	SF (%)	SA (%)
Gender	Male	54.55	38.99
	Female	45.45	61.01
Age	Young (age < 35)	35.23	35.85
	Middle age (35 ≤ age < 65)	38.64	45.91

Variables	Category	SF (%)	SA (%)
	Senior (age ≥ 65)	26.13	18.24
Race and ethnicity	White	38.64	29.56
	African American	4.55	11.95
	Asian	39.77	2.52
	Hispanic or Latino	12.50	55.97
	Other	4.54	—
Educational attainment	Level 1 (high school graduate or less, some college, vocational/technical training, or associate degree)	32.39	67.92
	Level 2 (bachelor’s degree)	38.07	20.13
	Level 3 (master’s degree, professional degree beyond bachelor's degree, e.g., M.D. and D.D.S., or doctoral degree, e.g., Ph.D.)	29.55	11.95
Employment	Full-time employed (paid)	45.59	40.25
	Part-time employed (paid)	10.23	13.84
	Self-employed	6.82	11.32
	Homemaker, unpaid volunteer or intern	3.98	8.18
	Retired	21.59	15.72
	Not currently employed	10.80	10.69
Household structure	Living alone	21.0	19.5
	Couple (without children)	29.5	22.0
	Nuclear family (i.e., couple living with child(ren))	19.3	20.1
	Shared adults	14.8	18.2
	Other	15.3	20.1
Household annual income	Low level (i.e., income < $75K)	42.61	74.21
	Medium level (i.e., $75K ≤ income < $150K)	32.95	22.01
	High level (i.e., income ≥ $150K)	24.44	3.78
Number of household vehicles	None	14.20	15.72
	1	40.91	42.14
	2	32.39	30.82
	3 or more	12.50	11.32
Holding alternative fuel vehicle	Yes	39.20	8.81
	No	60.80	91.19
Household residential ownership	Own/Buying (paying mortgage)	60.23	45.91
	Rent	34.09	42.14
	Other	5.68	11.95
Household residential duration	Less than 1 year	8.52	10.69
	1 to 5 years	27.84	40.88
	5 to 10 years	19.32	17.61

Variables	Category	SF (%)	SA (%)
	10 to 20 years	14.20	13.21
	20 years or more	30.11	17.61
Household residential type	Single-family house (detached house)	67.05	66.04
	Single-family house/townhouse (attached to one or more houses, each with separate entry) or multi-family house (3 or fewer apartments)	11.93	7.55
	Apartment	18.18	22.01
	Other	2.84	4.31
Sample size		**n = 176**	**n = 159**

4.2. Travel Behavior Profiles

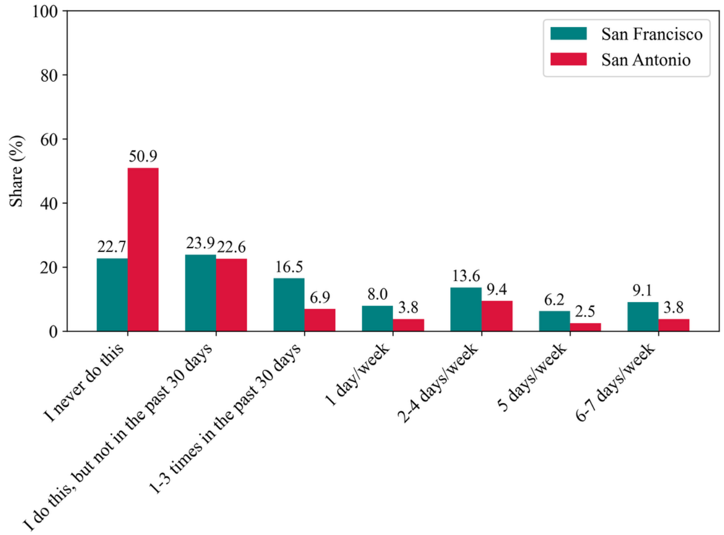
The statistical distribution of travel pattern across the two sample datasets is shown in Table 2. There are notable differences in travel-related characteristics between SF and SA. A higher share of SF respondents hold a driver's license (89.8%) compared to SA (79.3%), and a slightly larger share in SF are the primary driver of a household vehicle (84.1% versus 79.9%). Commute mode patterns also diverge. Driving alone is more common in SA (47.8%) than in SF (39.2%), while public transit use is significantly higher in SF (10.2%) compared to SA (2.5%), highlighting differences in transit availability or preferences. Active mode (e.g., bike and walk) share is marginally higher in SF (5.1%) than SA (4.4%). Commute distance distributions show SF has a larger proportion traveling 15 or more miles (24.4% versus 17.0%), whereas SA has more commuters in the 10 to 15-mile range (17.6% versus 11.4%). Both cities have similar shares of non-commuters or unemployed individuals (about 35%). These patterns suggest that SF residents rely more on public transit and active modes, whereas SA residents show greater dependence on driving alone and have a slightly more even distribution of commute distances under 15 miles.

Table 2. Distribution of Travel Pattern in San Francisco and San Antonio

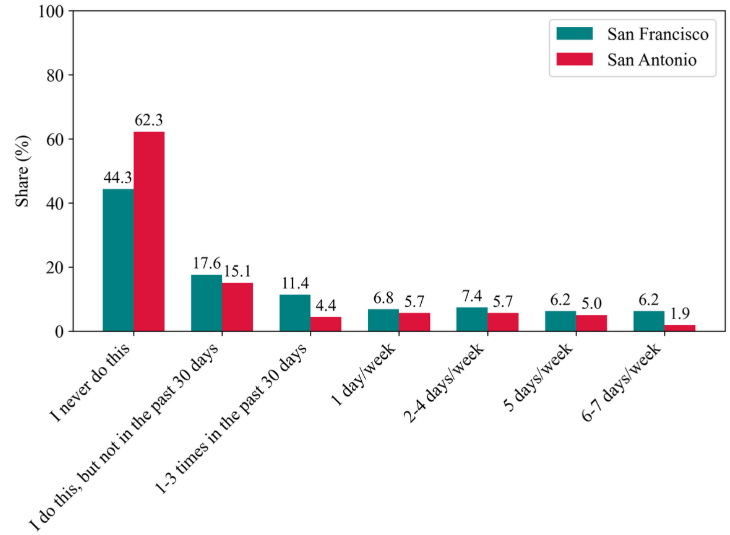
Variables	Category	SF (%)	SA (%)
Holding a driver's license	Yes	89.77	79.25
	No	10.23	20.75
Primary driver of at least a household vehicle	Yes	84.09	79.87
	No	15.91	20.13
Commute mode	Drive alone	39.20	47.80
	Carpool	6.82	6.92
	Bike or walk	5.11	4.40
	Public transit (bus or train)	10.23	2.52
	Other	2.84	3.77
	Not commuting or unemployed	35.80	34.59

Variables	Category	SF (%)	SA (%)
Commute distance (mile)	Less than 5	16.48	15.09
	5 to 10	11.93	15.72
	10 to 15	11.36	17.61
	15 or more	24.43	16.98
	Not commuting or unemployed	35.80	34.59
Sample size		**n = 176**	**n = 159**

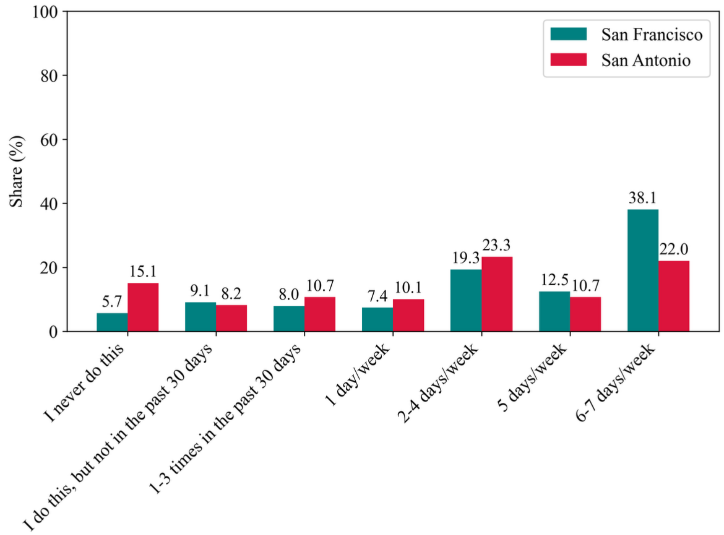
The distribution of travel frequency in Figure 2, reveals notable differences between SF and SA across all modes. For transit use, SF respondents report significantly higher regular usage, with lower shares in the "never" and "not in the past 30 days" categories, whereas SA shows much higher non-use. Biking and walking follow a similar trend. SF has smaller proportions of non-users and more frequent riders/walkers, particularly at moderate-to-high weekly frequencies, while SA shows heavier concentration in low or no usage. Carsharing is also more common in SF, with fewer respondents reporting "never" and higher shares using it at least occasionally. SA's participation is comparatively minimal. In contrast, rideshare services (both single and pooled) are more evenly split, though SF still shows more frequent use, SA retains a sizable proportion of occasional users. Overall, SF demonstrates a more multimodal travel pattern with greater integration of active and shared modes, while SA respondents are more concentrated in non-use or infrequent participation across most modes.



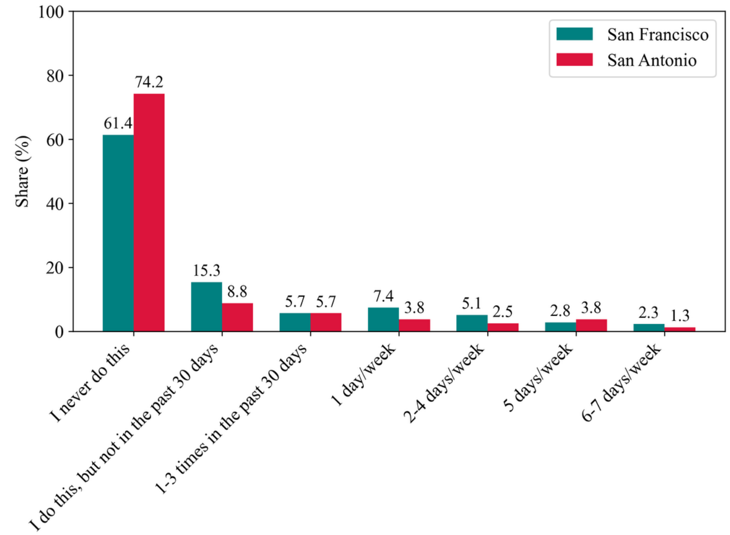
(a) Transit (Bus or Train)



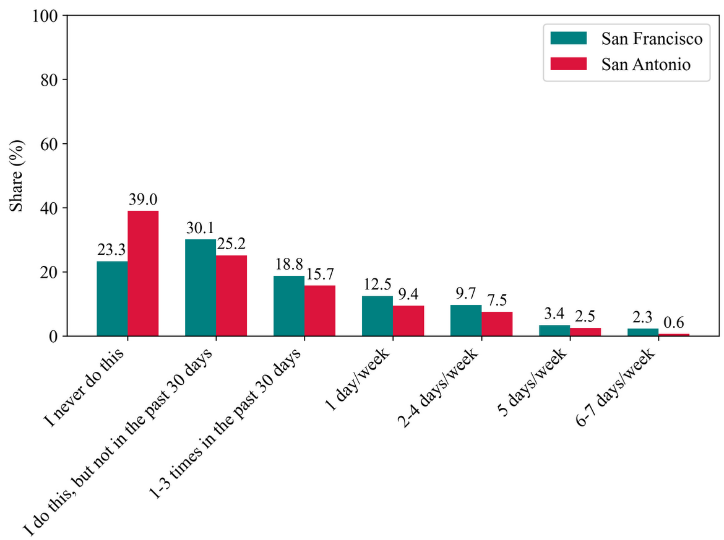
(b) Bike



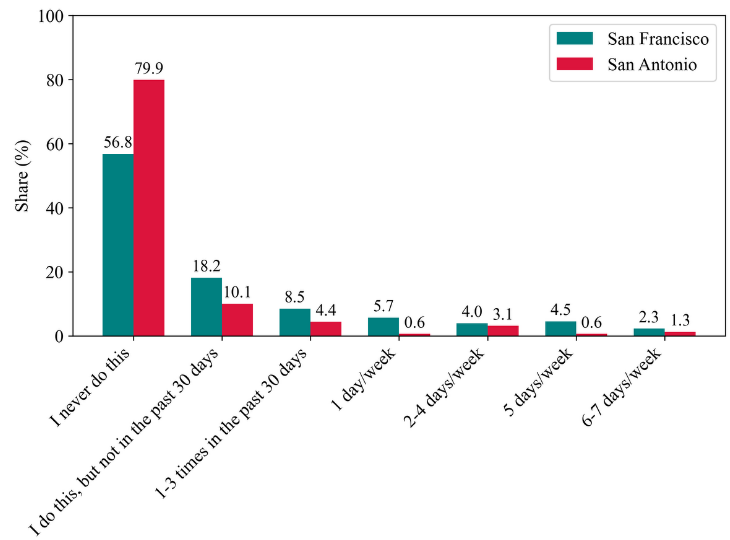
(c) Walk



(d) Carshare (such as Zipcar, Getaround, or Maven)



(e) Rideshare, single passenger (such as Uber or Lyft)



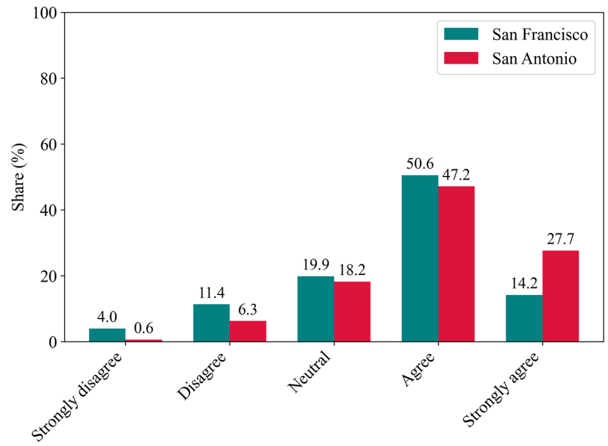
(f) Rideshare, carpool option (such as UberPOOL)

Figure 2. Distribution of Travel Mode Use in San Francisco and San Antonio

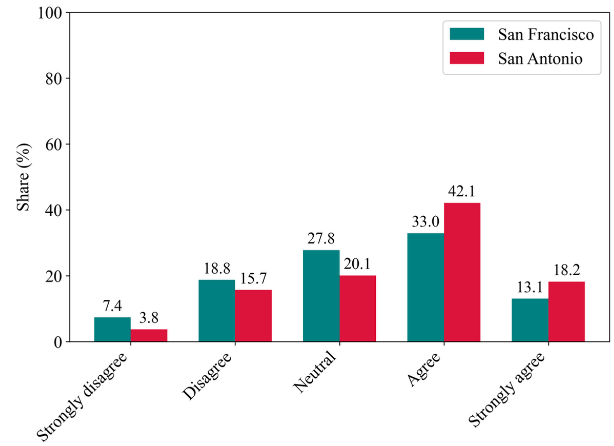
4.3. Attitudinal and Lifestyle Indicators

To better understand various dimensions of decision-making psychology, the study respondents were inquired about their opinions on a series of questions, which build two attitudinal and preferential constricts including emotionalism versus rationalism and utilitarianism versus deontology. The emotions versus rationality framework distinguishes between decisions primarily guided by feelings, intuition, and instinctive reactions versus those grounded in logical reasoning, deliberate evaluation, and evidence-based judgment. Emotion-driven decision-making can be advantageous in contexts requiring quick, adaptive responses, while rationality-driven approaches tend to dominate when accuracy, consistency, and analytical trade-off evaluations are needed. In contrast, the utilitarianism versus deontology dimension reflects other aspects. The former prioritizes outcomes that maximize overall well-being, even at the expense of individual rights or comfort, whereas the latter emphasize adherences to specific rules, duties, and principles regardless of the consequences. Overall, these dimensions provide a richer understanding of how individuals approach complex choices, blending cognitive style with specific orientation, and allowing for nuanced comparisons between populations or contexts such as the SF and SA samples in this study.

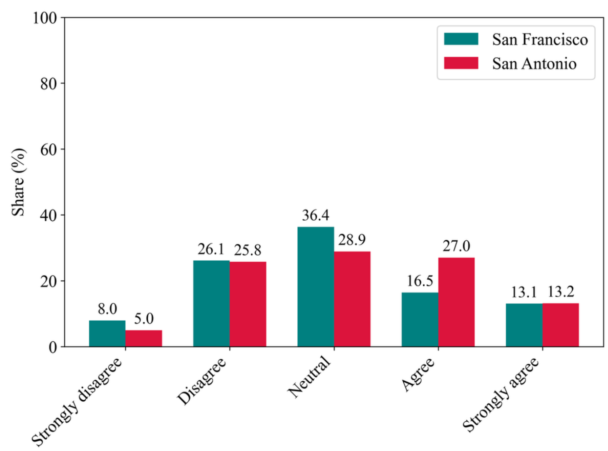
The distributions of the measurement indicators of these attitudes toward decision-making, each captured by the relevant question of the survey, along the emotions versus rationality spectrum reveal notable differences between the two city residents (Figure 3). For emotion-oriented tendencies, SF respondents show higher reliance on intuition and gut feelings in some measures compared to SA, whereas SA respondents are more likely to "agree" or "strongly agree" with making quick, instinct-based decisions. In contrast, rationality-oriented measures indicate generally high endorsement in both cities, with SA slightly ahead in deliberate, methodical decision-making and SF slightly ahead in weighing pros and cons. Both cities show strong alignment with logical reasoning, though SF leans marginally higher in strong agreement. Overall, SF respondents appear more balanced between emotional and rational cues, while SA respondents lean more toward rapid, instinctive decisions but also maintain strong rational decision-making endorsement. This suggests that interventions or policy framing may need to account for SF's openness to intuitive reasoning alongside rational deliberation, whereas in SA, quick emotional instincts coexist with strong rational frameworks.



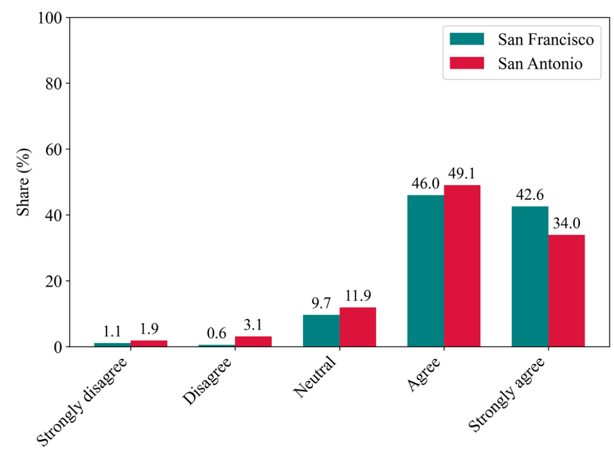
(a) Feelings and Intuition



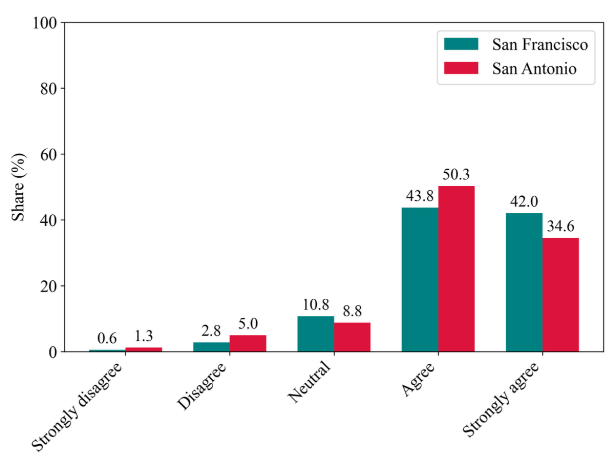
(b) Gut Feeling Even If Goes Against Data or Advice



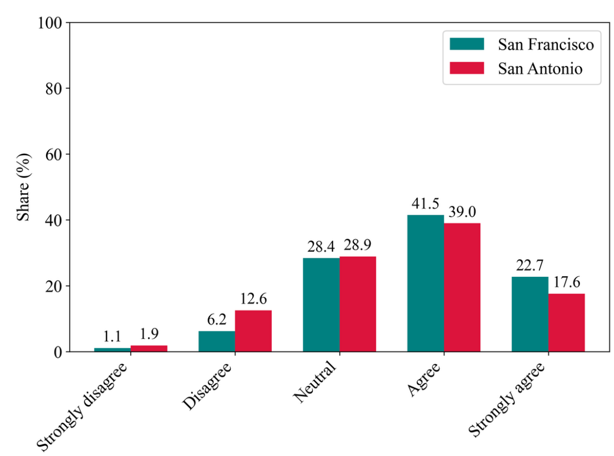
(c) Quickly, Based on Instincts or Emotions



(d) Logical Reasoning and Facts

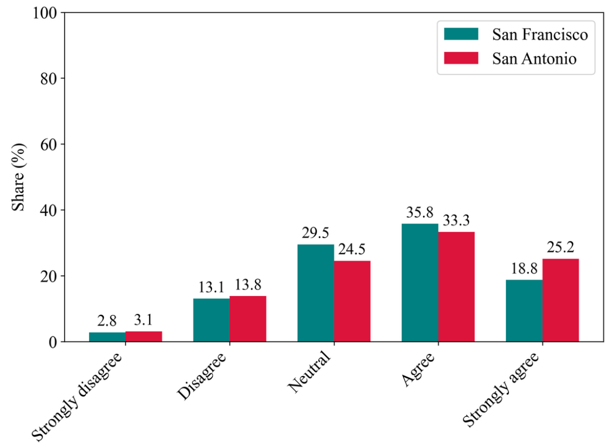


(e) Weighing All Pros and Cons Before Deciding

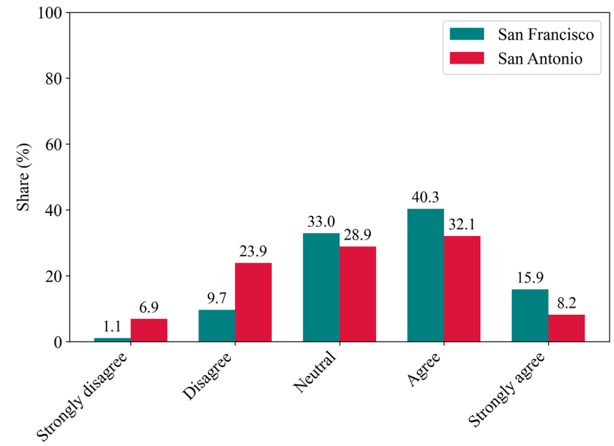


(f) Deliberately, Even If Takes Longer

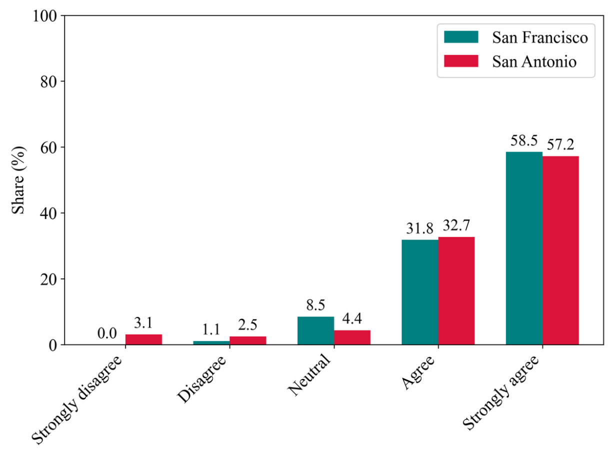
Figure 3. Distribution of Emotions-Driven vs. Rationality-Driven Decision-Making Criteria in San Francisco and San Antonio



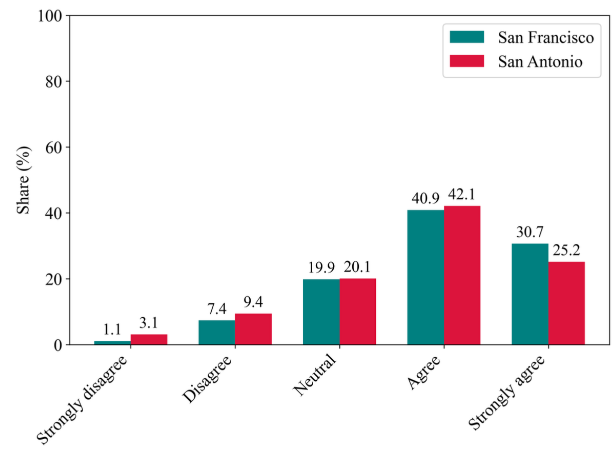
(a) I value individual freedom more than collective rules.



(b) I prefer making decisions that benefit most people, even if they harm a few.



(c) Protecting human life is more important than protecting property.



(d) I avoid situations where I have to choose between two bad outcomes.

Figure 4. Distribution of Utilitarian and Deontological Decision-Making Perspectives in San Francisco and San Antonio

The distributions of utilitarian and deontological decision-making perspectives in SF and SA (Figure 4) reveals both strong consensus on certain priorities and notable divergences between the two cities. Across both locations, there is significant agreement, particularly strong agreement, that protecting human life outweighs protecting property, indicating a shared deontological emphasis on intrinsic duties over material considerations (Figure 4(c)). Similarly, a majority in both cities express support for utilitarian reasoning, favoring decisions that maximize benefits for the greatest number even at the expense of a few, though SF shows a slightly higher share in strong agreement (Figure 4(b)). Divergence emerges in views on individual freedom versus collective rules, with SF respondents showing higher agreement and strong agreement, suggesting a stronger libertarian-leaning utilitarian orientation (Figure 4(a)). By contrast, avoiding complex dilemmas, which reflects a deontological preference, receives mixed responses, with SA residents more likely to agree or strongly agree with avoiding such situations, while SF respondents show higher neutrality or disagreement, possibly reflecting greater willingness to engage in complex trade-offs (Figure 4(d)).

Overall, these patterns suggest both cities value human life and societal benefit but differ in tolerance for complex situations and the balance between individual liberty and collective norms.

4.4. Autonomous Vehicle-Related Measures

4.4.1. AV Familiarity

Figure 5 highlights marked differences in AV familiarity between the two cities. In SA, lower familiarity levels dominate, with 23.3% of respondents reporting no awareness and 30.2% indicating no familiarity, which are substantially higher than the corresponding shares in the SF sample (2.3% and 13.1%, respectively). In contrast, SF respondents exhibit notably higher familiarity, with 38.1% reporting being very familiar without ride experience (compared to 12.6% in SA) and 26.1% being very familiar with ride experience (compared to merely 5.7% in SA). These differences suggest greater AV exposure in SF, in terms of both public awareness and direct experience (as is also reported by REF), likely reflecting a more mature AV presence, broader testing activities, and more extensive public discourse compared to SA.

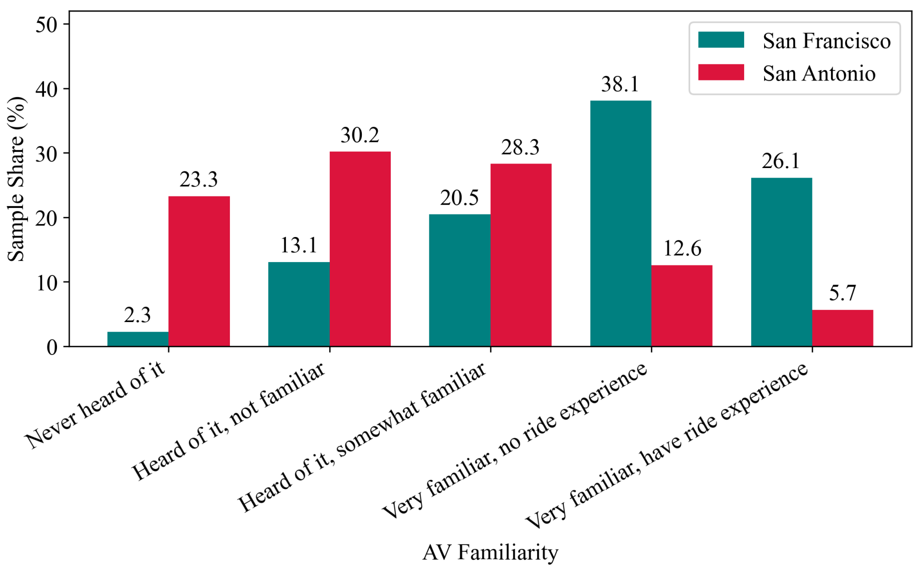
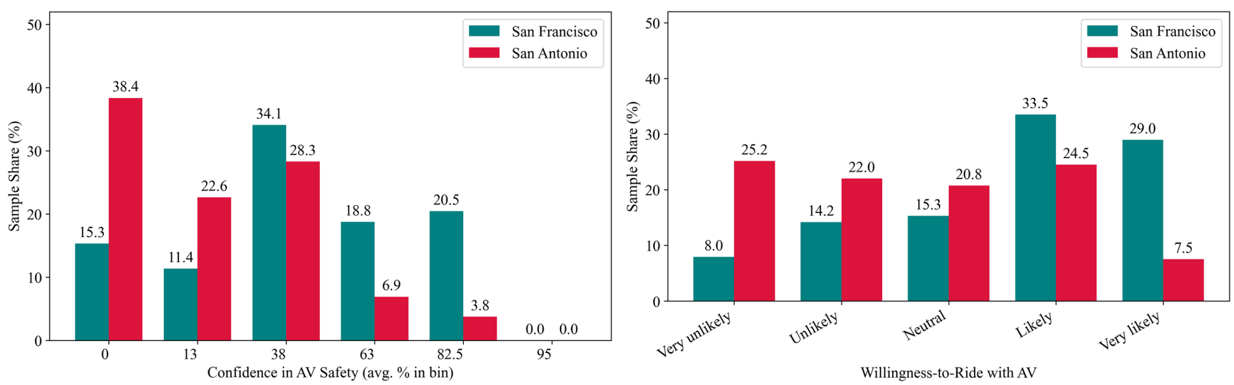


Figure 5. Distribution of Autonomous Vehicle Familiarity Levels in the San Francisco and San Antonio Sample Datasets

4.4.2. Baseline Attitudes toward AVs

The distribution of confidence in AV safety for the base scenario is shown in Figure 6(a). SA respondents are disproportionately concentrated at the lowest confidence level, with 38.4% reporting an average zero confidence compared to only 15.3% in SF. Similarly, 22.6% of SA participants fall into the 13% confidence bin, which is approximately twice as the SF share of 11.4%. In contrast, SF respondents dominate the higher confidence ranges with 20.5% reporting 82.5% confidence (versus 3.8% in SA) and 18.8% reporting 63% confidence (versus 6.9% in SA). Notably, the highest confidence category of 95% receives no responses in either city, suggesting a ceiling effect in how participants perceive AV safety under baseline conditions. These patterns indicate a statistically significant and practically meaningful gap, with SF residents on average exhibiting greater trust in AV safety than their SA counterparts.

Willingness-to-ride with AVs under the base scenario (Figure 6(b)) reveals a similarly pronounced city-level divide. In SF, responses are strongly skewed toward positive intentions with 33.5% reporting to be “Likely” and 29.0% “Very likely” riding in AV, compared to only 8.0% and 14.2%, respectively, in the two lowest-likelihood categories (“Very unlikely” and “Unlikely”). By contrast, SA presents a more cautious profile, with 25.2% “Very unlikely” and 22.0% “Unlikely,” and only 7.5% reporting “Very likely.” The proportion of “Neutral” responses is relatively comparable (15.3% in SF versus 20.8% in SA), suggesting similar levels of indecision among a smaller middle segment. Overall, the base scenario results indicate that SF respondents display substantially greater openness to riding in AV, while SA respondents exhibit stronger reluctance, aligning with the lower AV familiarity and safety confidence observed in Figure 5 and Figure 6(a).



(a) **(b)**

Figure 6. Distribution of Attitudes Toward Autonomous Vehicle in Base Scenario for San Francisco and San Antonio

4.4.3. Policy Scenario Attitudes toward AVs

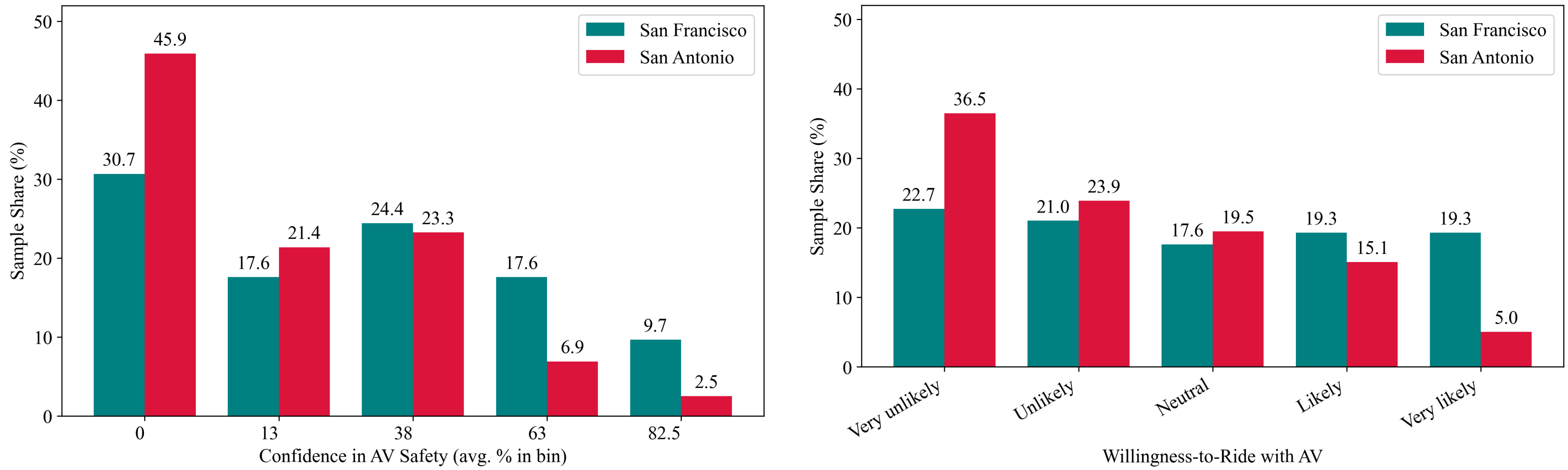
The distribution of attitudes toward AV for the four scenarios is shown in Figure 7. Across the four hypothetical scenarios, AV attitudes are observed with city-level differences yet with varying magnitudes depending on the scenario. In Scenario 1, where pedestrians are always prioritized over the on-board passengers, SA respondents show pronounced skepticism and hesitancy. Over 44% report zero confidence in AV safety compared to 22.2% in SF, while the highest-confidence share (82.5%) drops sharply for both cities. Willingness-to-ride follows the same polarizing pattern, with 36.5% of SA respondents “Very unlikely” to ride (versus 22.7% in SF) and only 5.0% “Very likely” (versus 19.3% in SF). Scenario 1 appears to polarize perceptions, amplifying low-confidence sentiment while eroding strong confidence, particularly in SF. This pattern suggests that an unconditional prioritization of pedestrians may heighten safety concerns among those worried about unpredictable AV behavior, even in a city with greater AV exposure.

Scenario 2 with prioritizing passengers only if pedestrians break laws moderates these extremes. In confidence, SF's high-confidence share rebounds modestly, while SA's zero-confidence share declines compared to Scenario 1. Willingness-to-ride also improves in both cities, with the "Very unlikely" share falling to 15.3% in SF and 33.3% in SA, and SF's "Very likely" share rising to 22.7%. These changes suggest that rule-based AV behavior may reassure respondents, especially in SF, by signaling predictability and alignment with traffic laws.

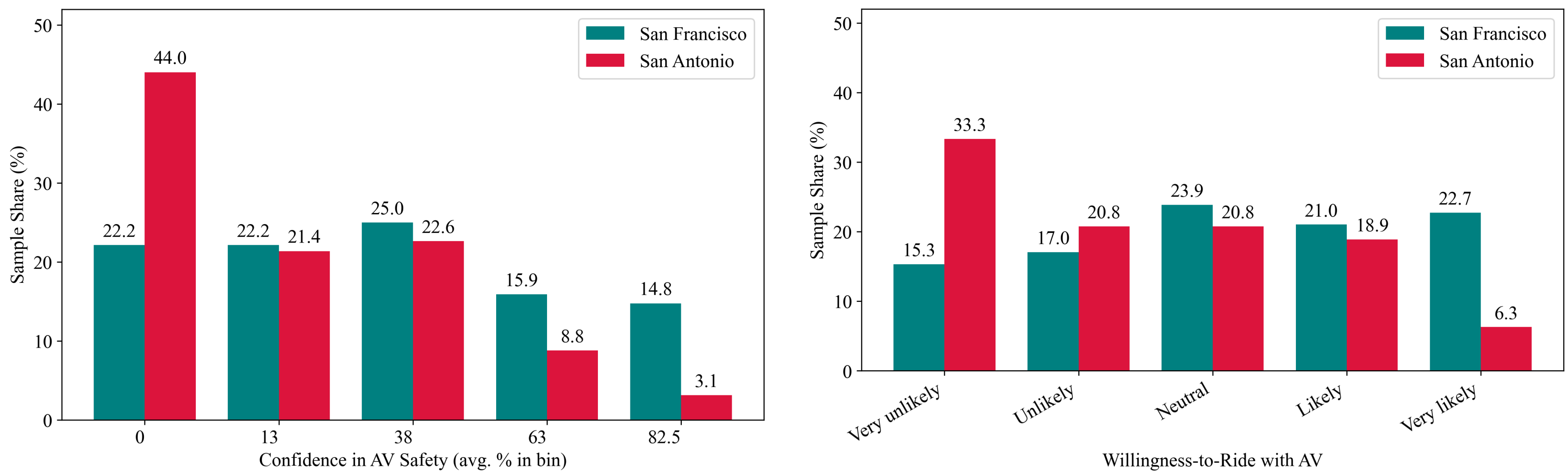
In Scenario 3 with the prioritization of children pedestrians, city-level gaps narrow further. Zero-confidence shares drop to 25.0% in SF and 40.3% in SA, and willingness-to-ride distributions converge slightly, producing the smallest gap across all scenarios. SA respondents show modest gains in mid- and high-likelihood categories, while SF maintains relatively balanced willingness levels. This alignment may reflect a shared consensus on prioritizing vulnerable road users, reducing perceived risk regardless of AV familiarity levels.

Finally, Scenario 4, which prioritized pedestrians if more numerous, brings mixed effects. In confidence, SA respondents return to their highest zero-confidence share (44.0%), while SF retains a relatively large high-confidence group, suggesting differing interpretations of "more numerous." Willingness-to-ride increases in SF, with 22.7% "Very likely," while SA records its highest "Very likely" share across all scenarios (20.1%). This rebound in SA indicates that some respondents may view numerical prioritization as a rational, data-driven safety rule, even if overall confidence remains lower than in SF.

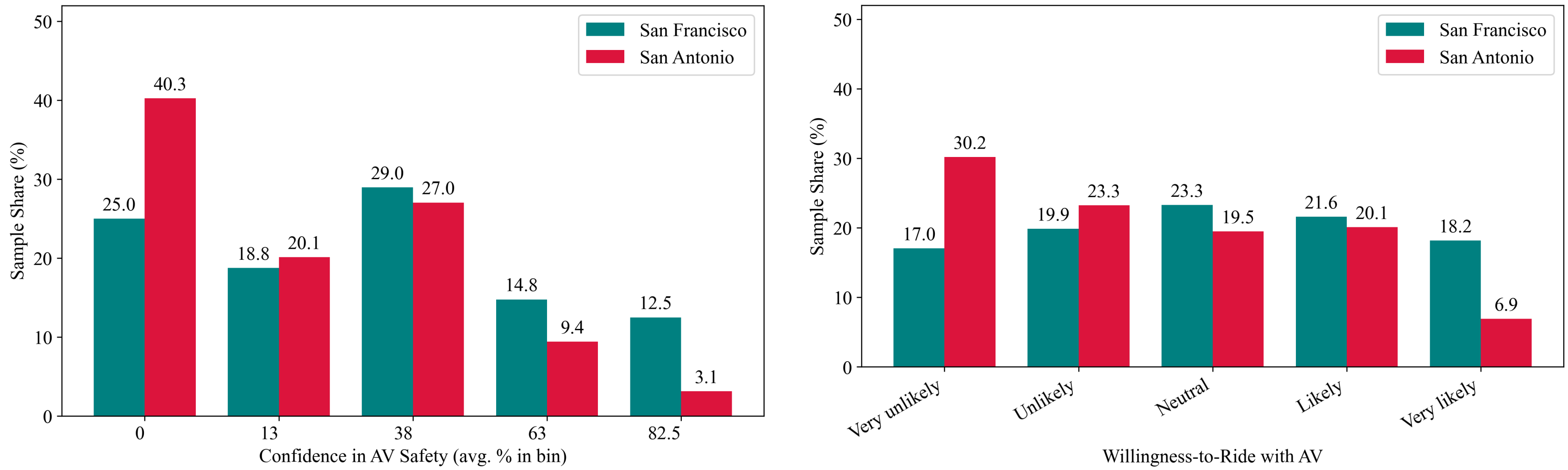
Overall, these patterns suggest that policies perceived as situationally justified (like Scenarios 2 and 4) can improve willingness-to-ride. However, rules seen as overly restrictive or biased (like Scenarios 1) may exacerbate hesitancy, particularly in lower-familiarity contexts like SA.



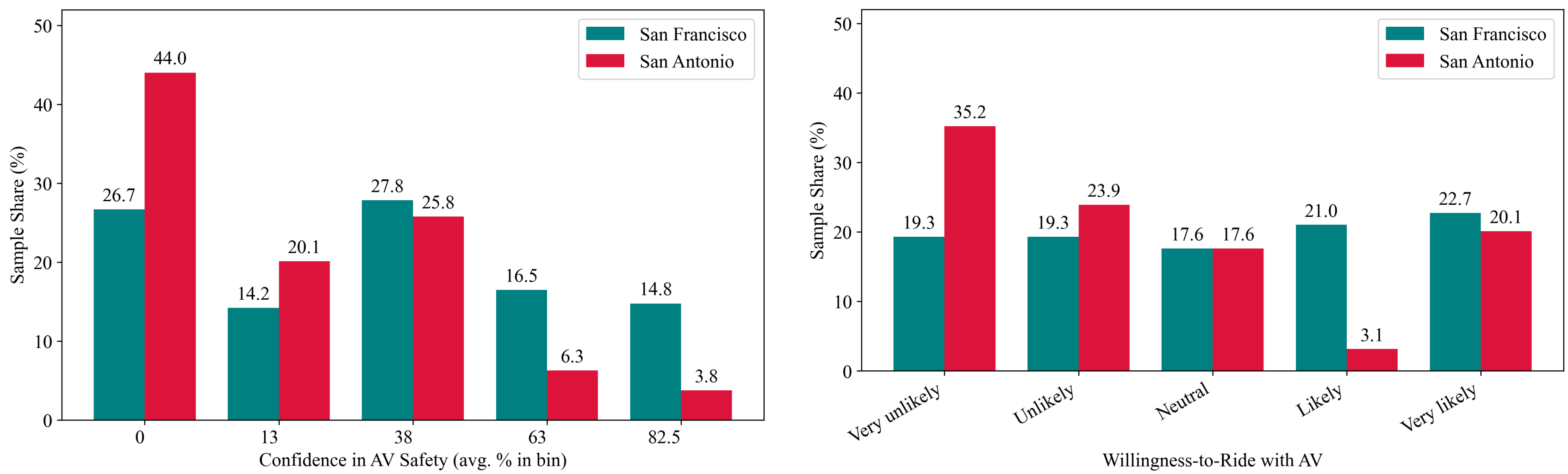
(b) Sc1 – Pedestrians Always Prioritized



(c) Sc2 – Passengers Prioritized If Pedestrian Breaks Laws



(d) Sc3 – Pedestrians Prioritized If Children



(e) Sc4 – Pedestrians Prioritized If More Numerous

Figure 7. Distribution of Attitudes Toward Autonomous Vehicle Across Four Hypothetical Scenarios in San Francisco and San Antonio

Table 3 reports statistical tests comparing SF and SA in confidence in AV safety and willingness-to-ride with AVs under the base and four hypothetical policy scenarios. The Chi-square (χ^2) test assesses differences in the overall response distribution, while MW test evaluates differences in central tendency (median responses). In the base scenario, both confidence in AV safety and willingness-to-ride differ sharply between cities, with both χ^2 and MW tests yielding p < 0.001, indicating very strong evidence of between-city differences. Across scenarios, this pattern largely persists. For confidence in AV safety, χ^2 results are significant (p < 0.01) in Scenarios 1, 2, and 4, but not in Scenario 3 (p = 0.104), whereas MW tests show highly significant differences (p < 0.001) for all scenarios, suggesting that median differences remain even when full distributions are more similar. For willingness-to-ride, both tests indicate significant differences (p ≤ 0.002) across all scenarios, with SF respondents consistently reporting higher willingness than SA respondents. Overall, city-level differences in AV perceptions are robust across policy contexts. The smallest confidence gap appears when AVs prioritize children pedestrians (Scenario 3), yet willingness-to-ride differences remain large in all scenarios, implying that adoption intentions are less sensitive to framing and more influenced by broader, location-specific attitudes toward AVs.

Table 3. Between-City Differences in Confidence in AV Safety and Willingness-to-Ride Across Policy Scenarios (Chi-Square and Mann–Whitney U Tests: San Francisco vs. San Antonio)

Scenario	Confidence in AV safety		Willingness-to-ride with AV	
	p(χ^2)	p(MW)	p(χ^2)	p(MW)
Base scenario	<0.001***	<0.001***	<0.001***	<0.001***
Sc1 – Pedestrians always prioritized	0.001**	<0.001***	<0.001***	<0.001***
Sc2 – Passengers prioritized if pedestrian breaks laws	0.003**	<0.001***	0.002**	<0.001***
Sc3 – Pedestrians prioritized if children	0.104	<0.001***	0.005**	<0.001***
Sc4 – Pedestrians prioritized if more numerous	<0.001***	<0.001***	<0.001***	<0.001***

Note: * p < 0.05; ** p < 0.01; *** p < 0.001. χ^2 = Chi-square test; MW = Mann–Whitney U test.

Table 4 summarizes within-city changes in confidence in AV safety and willingness-to-ride when moving from the baseline to each of the four policy scenarios. Median scores and MW U test p-values are reported for each comparison. In SF, both attitudes decline uniformly across scenarios, with median confidence and willingness-to-ride dropping from high (3.0) at baseline to neutral (2.0) in all cases. All changes are statistically significant at the 1% or 0.1% level, indicating that any departure from the baseline decision rule — whether prioritizing pedestrians or passengers — substantially erodes positive AV perceptions. SA begins with lower baseline medians (2.0) and shows smaller, mostly non-significant shifts. Significant declines occur only in Scenario 1 for willingness-to-ride and in Scenarios 1 and 2 for confidence, and even these changes are modest (–1.0). Overall, SF respondents display greater sensitivity to AV decision rules, while SA respondents' attitudes remain comparatively stable regardless of scenario framing.

Table 4. Within-City Changes in Confidence in AV Safety and Willingness-to-Ride Across Policy Scenarios (Median Level, Change from Baseline, and Mann–Whitney U Tests)

Scenario	Confidence in AV safety			Willingness-to-ride with AV		
	Median	Δ	p(MW)	Median	Δ	p(MW)
San Francisco						
Base scenario	3.0	—	—	3.0	—	—
Sc1 – Pedestrians always prioritized	2.0	-1.0	<0.001***	2.0	-1.0	<0.001***
Sc2 – Passengers prioritized if pedestrian breaks laws	2.0	-1.0	0.007**	2.0	-1.0	0.003**
Sc3 – Pedestrians prioritized if children	2.0	-1.0	<0.001***	2.0	-1.0	<0.001***
Sc4 – Pedestrians prioritized if more numerous	2.0	-1.0	0.002**	2.0	-1.0	<0.001***
San Antonio						
Base scenario	2.0	—	—	2.0	—	—
Sc1 – Pedestrians always prioritized	1.0	-1.0	0.121	1.0	-1.0	0.006**
Sc2 – Passengers prioritized if pedestrian breaks laws	1.0	-1.0	0.768	1.0	-1.0	0.101
Sc3 – Pedestrians prioritized if children	1.0	-1.0	0.604	1.0	-1.0	0.232
Sc4 – Pedestrians prioritized if more numerous	1.0	-1.0	0.229	1.0	-1.0	0.120

Notes: Confidence in AV safety is measured on 5-point ordinal scales: 0 = Very low, 1 = Low, 2 = Neutral, 3 = High, 4 = Very high; Willingness-to-ride with AV is measured on 5-point ordinal scales: 0 = Very unlikely, 1 = Unlikely, 2 = Neutral, 3 = Likely, 4 = Very likely. Δ = difference in median from the baseline scenario (negative values indicate lower ratings). p(MW) = p-value from the Mann–Whitney U test comparing scenario ratings with the baseline within each city. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5. Results

This section presents the simulation outcomes generated from DBN model, focusing on how alternative policy scenarios influence public attitudes toward AV safety and willingness-to-ride. The results are organized to first provide an overall view of predicted changes under each scenario, then to examine heterogeneity by city and baseline attitude strata, enabling a nuanced understanding of policy responsiveness.

The policy analysis uses the DBN trained on respondents' baseline attitudes and their stated preferences under four hypothetical policies governing pedestrian–passenger prioritization in crash-imminent situations. In the simulations, baseline confidence and willingness-to-ride levels, together with scenario-specific policy codes, determine the post-policy probability distribution of attitudes. The first approach, which is empirical-mix, weights the model's scenario-level predictions by the observed distribution of baseline attitudes in each city, producing expected values (E[level] and E[mid%]) that reflect realistic population heterogeneity. Differences from the base scenario (Δ) are calculated within each city, and MW U tests assess whether scenario-induced changes are statistically significant. This approach allows direct comparison of simulated policy impacts across SF and SA, accounting for baseline variation and enabling policy-relevant interpretation of directional and magnitude changes.

Table 5 summarizes the simulated confidence in AV safety under the baseline and four policy scenarios for SF and SA, based on the empirical-mix approach. At the baseline, expected confidence is higher in SF (E[level] = 2.636, or 62.7%) than in SA (E[level] = 1.484, or 39.7%). In SF, all four policy scenarios produced statistically significant declines in confidence relative to baseline, with the largest decreases for scenarios prioritizing pedestrians when children are present (–0.632 level; –12.6%) and when pedestrians are more numerous (–0.634 level; –12.7%). By contrast, in SA, none of the changes from baseline are statistically significant. Small negative shifts occur for Scenario 1 (–0.1 level; –2.0%), while Scenario 3 and Scenario 4 show moderate positive changes (+0.502 and +0.508 levels; about +10%), though without statistical significance. These results indicate that SF respondents' simulated confidence is more sensitive to changes in pedestrian–passenger prioritization policies, whereas SA respondents' confidence remained comparatively stable across scenarios.

Table 5. Empirical-Mix Simulated Confidence in AV Safety by Policy Scenario for San Francisco and San Antonio

Scenario	E[level]	E[mid%]	Δ[level]	Δ[mid%]	p(MW)
San Francisco					
Base scenario	2.636	62.7	—	—	—
Sc1 – Pedestrians always prioritized	2.083	51.7	–0.553	–11.1	<0.001***
Sc2 – Passengers prioritized if pedestrian breaks laws	2.244	54.9	–0.393	–7.9	0.007**
Sc3 – Pedestrians prioritized if children	2.005	50.1	–0.632	–12.6	<0.001***
Sc4 – Pedestrians prioritized if more numerous	2.002	50.0	–0.634	–12.7	0.002**
San Antonio					
Base scenario	1.484	39.7	—	—	—
Sc1 – Pedestrians always prioritized	1.384	37.7	–0.1	–2.0	0.121
Sc2 – Passengers prioritized if pedestrian breaks laws	1.512	40.2	0.028	0.6	0.768
Sc3 – Pedestrians prioritized if children	1.987	49.7	0.502	10.0	0.604
Sc4 – Pedestrians prioritized if more numerous	1.993	49.9	0.508	10.2	0.229

Note: E[level] = expected confidence in AV safety on 0–4 scale; E[mid%] = midpoint-scaled equivalent (10–90%). Δ = difference from base scenario within the same city. p(MW) from Mann–Whitney U tests versus base; significance: * p<0.05, ** p<0.01, *** p<0.001. Negative Δ = lower confidence; positive Δ = higher confidence.

Table 6 presents the simulated willingness-to-ride with AVs under the baseline and four policy scenarios for SF and SA. At baseline, SF respondents show higher willingness (E[level] = 2.614) than SA respondents (E[level] = 1.673). In SF, all four policy scenarios produce statistically significant declines in willingness-to-ride, with the largest decrease occurring when pedestrians are always prioritized (–0.702 level), followed by smaller yet still significant declines across the other three scenarios (–0.593 to –0.603 levels). In contrast, SA shows small positive changes from baseline under all scenarios, with the largest gains for Scenario 2 (+0.222 level) and Scenarios 3 and 4 (+0.200 levels), though only Scenario 1 (+0.113) reaches statistical significance. These patterns suggest that SF respondents' willingness-to-ride is highly sensitive to changes in pedestrian–passenger prioritization, consistently decreasing under all scenarios, whereas SA respondents tend to show modest increases, particularly when policies favor passenger rights or balance between groups.

Table 6. Empirical-Mix Simulated Willingness-to-Ride with AVs by Policy Scenario for San Francisco and San Antonio

Scenario	E[level]	E[mid%]	Δ[level]	Δ[mid%]	p(MW)
San Francisco					
Base scenario	2.614	3.614	—	—	—
Sc1 – Pedestrians always prioritized	1.912	2.912	–0.702	–0.702	<0.001***
Sc2 – Passengers prioritized if pedestrian breaks laws	2.010	3.010	–0.603	–0.603	0.003**
Sc3 – Pedestrians prioritized if children	2.020	3.020	–0.593	–0.593	<0.001***
Sc4 – Pedestrians prioritized if more numerous	2.020	3.020	–0.593	–0.593	<0.001***
San Antonio					
Base scenario	1.673	2.673	—	—	—
Sc1 – Pedestrians always prioritized	1.785	2.785	0.113	0.113	0.006**
Sc2 – Passengers prioritized if pedestrian breaks laws	1.895	2.895	0.222	0.222	0.101
Sc3 – Pedestrians prioritized if children	1.873	2.873	0.200	0.200	0.232
Sc4 – Pedestrians prioritized if more numerous	1.873	2.873	0.200	0.200	0.120

Note: E[level] = expected willingness-to-ride with AV on 0–4 scale; E[mid%] = midpoint-scaled equivalent (10–90%). Δ = difference from base scenario within the same city. p(MW) from Mann–Whitney U tests versus base; significance: * p<0.05, ** p<0.01, *** p<0.001. Negative Δ = lower confidence; positive Δ = higher confidence.

In the second stage of the policy simulation framework, we stratified results by baseline confidence levels to capture heterogeneity in respondents' simulated reactions to the four policy scenarios. Table 7-Table 8 and Table 9-Table 10 present the outcomes respectively for confidence in AV safety and willingness-to-ride, with the simulations applying each policy's estimated effects to baseline distributions to generate expected values and probability profiles for each group. Across both cities, baseline confidence in AV safety strongly shapes simulated policy responses for both confidence and willingness-to-ride outcomes. In SF and SA, Scenarios 1 and 2 amplify pre-existing differences — low-confidence groups remain at the bottom of the scale, while high-confidence groups sustain much higher values — indicating that these policies polarize perceptions. In contrast, Scenarios 3 and 4 produce a convergence effect, pulling all baseline groups toward midscale values and reducing heterogeneity. For willingness-to-ride, high-confidence groups in both cities consistently score higher and show more balanced distributions across higher willingness categories, whereas low-confidence groups cluster in the lowest categories. Mid-confidence respondents respond most to Scenario 2 in both contexts. While patterns are similar across cities, policy impacts appear slightly more pronounced in SF, particularly in elevating mid- and high-confidence respondents' willingness-to-ride.

Table 7. Scenario-Based Simulations of Confidence in AV Safety for San Francisco, Stratified by Baseline Confidence Levels

Level	Weight (%)	E[Level]	E[mid%]	P(Level=0)	P(Level=1)	P(Level=2)	P(Level=3)	P(Level=4)
Sc1 – Pedestrians always prioritized								
0	9.1	0.435	18.697	0.79	0.056	0.111	0.014	0.028
1	11.9	0.907	28.137	0.499	0.25	0.116	0.116	0.02
2	20.5	1.847	46.936	0.197	0.197	0.298	0.175	0.132
3	23.3	2.328	56.553	0.129	0.143	0.172	0.385	0.172
4	35.2	2.882	67.644	0.141	0.077	0.09	0.141	0.55
Low (0–1)	21	0.703	24.055	0.625	0.166	0.114	0.072	0.023
Mid (2)	20.5	1.847	46.936	0.197	0.197	0.298	0.175	0.132
High (3–4)	58.5	2.661	63.229	0.136	0.103	0.122	0.238	0.4
Sc2 – Passengers prioritized if pedestrian breaks laws								
0	9.1	0.449	18.974	0.762	0.098	0.098	0.014	0.028
1	11.9	1.022	30.438	0.403	0.288	0.231	0.039	0.039
2	20.5	1.998	49.966	0.016	0.016	0.938	0.017	0.014
3	23.3	2.684	63.677	0.072	0.072	0.157	0.499	0.2
4	35.2	2.972	69.434	0.103	0.052	0.18	0.103	0.563
Low (0–1)	21	0.774	25.481	0.559	0.206	0.173	0.028	0.034
Mid (2)	20.5	1.998	49.966	0.016	0.016	0.938	0.017	0.014
High (3–4)	58.5	2.857	67.142	0.09	0.06	0.171	0.261	0.419
Sc3 – Pedestrians prioritized if children								
0	9.1	1.961	49.217	0.213	0.198	0.197	0.196	0.195
1	11.9	1.978	49.57	0.203	0.205	0.198	0.197	0.197
2	20.5	1.996	49.92	0.197	0.2	0.208	0.199	0.195
3	23.3	2.015	50.292	0.196	0.197	0.201	0.208	0.198
4	35.2	2.023	50.462	0.198	0.196	0.198	0.199	0.208
Low (0–1)	21	1.971	49.417	0.208	0.202	0.198	0.196	0.196
Mid (2)	20.5	1.996	49.92	0.197	0.2	0.208	0.199	0.195
High (3–4)	58.5	2.02	50.394	0.198	0.196	0.199	0.203	0.204
Sc4 – Pedestrians prioritized if more numerous								
0	9.1	1.98	49.603	0.207	0.199	0.198	0.198	0.198
1	11.9	1.99	49.797	0.202	0.201	0.2	0.198	0.198
2	20.5	1.996	49.916	0.199	0.201	0.204	0.2	0.197
3	23.3	2.007	50.137	0.199	0.198	0.2	0.203	0.2
4	35.2	2.013	50.252	0.199	0.198	0.198	0.2	0.205
Low (0–1)	21	1.986	49.713	0.204	0.2	0.199	0.198	0.198
Mid (2)	20.5	1.996	49.916	0.199	0.201	0.204	0.2	0.197
High (3–4)	58.5	2.01	50.207	0.199	0.198	0.199	0.201	0.203

Table 8. Scenario-Based Simulations of Confidence in AV Safety for San Antonio, Stratified by Baseline Confidence Levels

Level	Weight (%)	E[Level]	E[mid%]	P(Level=0)	P(Level=1)	P(Level=2)	P(Level=3)	P(Level=4)
Sc1 – Pedestrians always prioritized								
0	32.1	0.435	18.697	0.79	0.056	0.111	0.014	0.028
1	16.4	0.907	28.137	0.499	0.25	0.116	0.116	0.02
2	29.6	1.847	46.936	0.197	0.197	0.298	0.175	0.132
3	15.1	2.328	56.553	0.129	0.143	0.172	0.385	0.172
4	6.9	2.882	67.644	0.141	0.077	0.09	0.141	0.55
Low (0–1)	48.4	0.594	21.885	0.692	0.121	0.113	0.049	0.025
Mid (2)	29.6	1.847	46.936	0.197	0.197	0.298	0.175	0.132
High (3–4)	22	2.502	60.039	0.133	0.122	0.146	0.309	0.291
Sc2 – Passengers prioritized if pedestrian breaks laws								
0	32.1	0.449	18.974	0.762	0.098	0.098	0.014	0.028
1	16.4	1.022	30.438	0.403	0.288	0.231	0.039	0.039
2	29.6	1.998	49.966	0.016	0.016	0.938	0.017	0.014
3	15.1	2.684	63.677	0.072	0.072	0.157	0.499	0.2
4	6.9	2.972	69.434	0.103	0.052	0.18	0.103	0.563
Low (0–1)	48.4	0.642	22.845	0.641	0.162	0.142	0.023	0.032
Mid (2)	29.6	1.998	49.966	0.016	0.016	0.938	0.017	0.014
High (3–4)	22	2.774	65.486	0.082	0.065	0.164	0.375	0.314
Sc3 – Pedestrians prioritized if children								
0	32.1	1.961	49.217	0.213	0.198	0.197	0.196	0.195
1	16.4	1.978	49.57	0.203	0.205	0.198	0.197	0.197
2	29.6	1.996	49.92	0.197	0.2	0.208	0.199	0.195
3	15.1	2.015	50.292	0.196	0.197	0.201	0.208	0.198
4	6.9	2.023	50.462	0.198	0.196	0.198	0.199	0.208
Low (0–1)	48.4	1.967	49.336	0.21	0.201	0.198	0.196	0.196
Mid (2)	29.6	1.996	49.92	0.197	0.2	0.208	0.199	0.195
High (3–4)	22	2.017	50.345	0.197	0.197	0.2	0.205	0.201
Sc4 – Pedestrians prioritized if more numerous								
0	32.1	1.98	49.603	0.207	0.199	0.198	0.198	0.198
1	16.4	1.99	49.797	0.202	0.201	0.2	0.198	0.198
2	29.6	1.996	49.916	0.199	0.201	0.204	0.2	0.197
3	15.1	2.007	50.137	0.199	0.198	0.2	0.203	0.2
4	6.9	2.013	50.252	0.199	0.198	0.198	0.2	0.205
Low (0–1)	48.4	1.983	49.669	0.206	0.2	0.199	0.198	0.198
Mid (2)	29.6	1.996	49.916	0.199	0.201	0.204	0.2	0.197
High (3–4)	22	2.009	50.173	0.199	0.198	0.2	0.202	0.201

Table 9. Scenario-Based Simulations of Willingness-to-Ride with AV for San Francisco, Stratified by Baseline Willingness Levels

Level	Weight (%)	E[Level]	E[mid%]	P(Level=0)	P(Level=1)	P(Level=2)	P(Level=3)	P(Level=4)
Sc1 – Pedestrians always prioritized								
0	8	1.585	2.585	0.347	0.18	0.165	0.157	0.151
1	14.2	1.618	2.618	0.244	0.304	0.175	0.144	0.133
2	15.3	1.912	2.912	0.092	0.1	0.673	0.073	0.062
3	33.5	1.93	2.93	0.176	0.214	0.239	0.246	0.125
4	29	2.124	3.124	0.19	0.19	0.176	0.194	0.25
Low (0–1)	22.2	1.606	2.606	0.281	0.26	0.171	0.149	0.139
Mid (2)	15.3	1.912	2.912	0.092	0.1	0.673	0.073	0.062
High (3–4)	62.5	2.02	3.02	0.182	0.203	0.21	0.222	0.183
Sc2 – Passengers prioritized if pedestrian breaks laws								
0	8	1.597	2.597	0.33	0.207	0.144	0.174	0.145
1	14.2	1.827	2.827	0.149	0.302	0.256	0.161	0.133
2	15.3	1.989	2.989	0.02	0.022	0.922	0.021	0.016
3	33.5	2.118	3.118	0.136	0.137	0.3	0.325	0.101
4	29	2.1	3.1	0.156	0.21	0.214	0.217	0.203
Low (0–1)	22.2	1.745	2.745	0.214	0.268	0.216	0.165	0.137
Mid (2)	15.3	1.989	2.989	0.02	0.022	0.922	0.021	0.016
High (3–4)	62.5	2.11	3.11	0.145	0.171	0.26	0.275	0.148
Sc3 – Pedestrians prioritized if children								
0	8	1.632	2.632	0.336	0.184	0.154	0.162	0.164
1	14.2	1.702	2.702	0.203	0.303	0.219	0.141	0.134
2	15.3	1.959	2.959	0.086	0.094	0.66	0.092	0.067
3	33.5	2.092	3.092	0.162	0.179	0.222	0.278	0.159
4	29	2.234	3.234	0.174	0.166	0.187	0.2	0.274
Low (0–1)	22.2	1.677	2.677	0.251	0.26	0.196	0.149	0.145
Mid (2)	15.3	1.959	2.959	0.086	0.094	0.66	0.092	0.067
High (3–4)	62.5	2.157	3.157	0.168	0.173	0.205	0.242	0.212
Sc4 – Pedestrians prioritized if more numerous								
0	8	1.632	2.632	0.336	0.184	0.154	0.162	0.164
1	14.2	1.701	2.701	0.203	0.303	0.218	0.141	0.134
2	15.3	1.958	2.958	0.087	0.095	0.659	0.092	0.067
3	33.5	2.091	3.091	0.163	0.179	0.221	0.278	0.159
4	29	2.235	3.235	0.174	0.165	0.187	0.2	0.274
Low (0–1)	22.2	1.676	2.676	0.251	0.26	0.195	0.149	0.145
Mid (2)	15.3	1.958	2.958	0.087	0.095	0.659	0.092	0.067
High (3–4)	62.5	2.158	3.158	0.168	0.173	0.205	0.242	0.213

Table 10. Scenario-Based Simulations of Willingness-to-Ride with AV for San Antonio, Stratified by Baseline Willingness Levels

Level	Weight (%)	E[Level]	E[mid%]	P(Level=0)	P(Level=1)	P(Level=2)	P(Level=3)	P(Level=4)
Sc1 – Pedestrians always prioritized								
0	25.2	1.585	2.585	0.347	0.18	0.165	0.157	0.151
1	22	1.618	2.618	0.244	0.304	0.175	0.144	0.133
2	20.8	1.912	2.912	0.092	0.1	0.673	0.073	0.062
3	24.5	1.93	2.93	0.176	0.214	0.239	0.246	0.125
4	7.5	2.124	3.124	0.19	0.19	0.176	0.194	0.25
Low (0–1)	47.2	1.6	2.6	0.299	0.238	0.169	0.151	0.143
Mid (2)	20.8	1.912	2.912	0.092	0.1	0.673	0.073	0.062
High (3–4)	32.1	1.976	2.976	0.179	0.208	0.224	0.234	0.154
Sc2 – Passengers prioritized if pedestrian breaks laws								
0	25.2	1.597	2.597	0.33	0.207	0.144	0.174	0.145
1	22	1.827	2.827	0.149	0.302	0.256	0.161	0.133
2	20.8	1.989	2.989	0.02	0.022	0.922	0.021	0.016
3	24.5	2.118	3.118	0.136	0.137	0.3	0.325	0.101
4	7.5	2.1	3.1	0.156	0.21	0.214	0.217	0.203
Low (0–1)	47.2	1.705	2.705	0.245	0.251	0.196	0.168	0.14
Mid (2)	20.8	1.989	2.989	0.02	0.022	0.922	0.021	0.016
High (3–4)	32.1	2.114	3.114	0.141	0.155	0.28	0.3	0.125
Sc3 – Pedestrians prioritized if children								
0	25.2	1.632	2.632	0.336	0.184	0.154	0.162	0.164
1	22	1.702	2.702	0.203	0.303	0.219	0.141	0.134
2	20.8	1.959	2.959	0.086	0.094	0.66	0.092	0.067
3	24.5	2.092	3.092	0.162	0.179	0.222	0.278	0.159
4	7.5	2.234	3.234	0.174	0.166	0.187	0.2	0.274
Low (0–1)	47.2	1.665	2.665	0.274	0.239	0.184	0.152	0.15
Mid (2)	20.8	1.959	2.959	0.086	0.094	0.66	0.092	0.067
High (3–4)	32.1	2.125	3.125	0.165	0.176	0.213	0.26	0.186
Sc4 – Pedestrians prioritized if more numerous								
0	25.2	1.632	2.632	0.336	0.184	0.154	0.162	0.164
1	22	1.701	2.701	0.203	0.303	0.218	0.141	0.134
2	20.8	1.958	2.958	0.087	0.095	0.659	0.092	0.067
3	24.5	2.091	3.091	0.163	0.179	0.221	0.278	0.159
4	7.5	2.235	3.235	0.174	0.165	0.187	0.2	0.274
Low (0–1)	47.2	1.664	2.664	0.274	0.239	0.184	0.152	0.15
Mid (2)	20.8	1.958	2.958	0.087	0.095	0.659	0.092	0.067
High (3–4)	32.1	2.125	3.125	0.165	0.176	0.213	0.259	0.186

6. Concluding Remarks

This study develops and applies a Dynamic Bayesian Network (DBN) framework to model changes in public confidence in AV safety and willingness-to-ride under hypothetical pedestrian–passenger prioritization policies. Using baseline survey data from San Francisco (SF) and San Antonio (SA), the model incorporates both intra-slice and inter-slice dependencies, enabling scenario-based simulations and counterfactual analyses. The empirical-mix results show that SF respondents generally begin with higher baseline confidence and willingness-to-ride, yet are more sensitive to policy changes, exhibiting significant declines when policies strongly prioritized pedestrians. In contrast, SA respondents display comparatively stable or slightly positive responses, especially under policies that prioritized passengers and balanced priorities. Stratifying results by baseline confidence revealed that some policies amplify pre-existing differences, while others promote convergence toward midscale attitudes, highlighting heterogeneous policy impacts across confidence segments.

While the DBN approach offers valuable insights into causal and temporal dynamics in AV policy acceptance, the present study has several limitations. The analysis relies on stated preferences datasets, which may not fully reflect actual behavior in real-world settings. The temporal dimension is limited to two time slices, restricting the capture of longer-term attitude evolution. Moreover, the policy scenarios examined are specific to pedestrian–passenger prioritization in crash-imminent contexts, which may limit generalizability to other AV policy domains. Future research should integrate revealed preferences and longitudinal panel data, expand to a broader set of policy interventions, and explore richer temporal structures in DBNs to capture multi-stage adoption processes. Incorporating additional contextual variables, such as media exposure, trust in institutions, and prior AV experience, could further enhance the explanatory power and policy relevance of the framework.

References

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

2. Bergmann, L.T., Schlicht, L., Meixner, C., König, P., Pipa, G., Boshammer, S., Stephan, A., 2018. Autonomous vehicles require socio-political acceptance—An empirical and philosophical perspective on the problem of moral decision making. Frontiers in Behavioral Neuroscience 12, 31. https://doi.org/10.3389/fnbeh.2018.00031

3. Dagum, P., Galper, A., Horvitz, E., 1992. Dynamic network models for forecasting, in: Uncertainty in Artificial Intelligence. Elsevier, pp. 41–48. https://doi.org/10.1016/B978-1-4832-8287-9.50010-4

4. Friedman, N., Geiger, D., Goldszmidt, M., 1997. Bayesian network classifiers. Machine Learning 29, 131–163. https://doi.org/10.1023/A:1007465528199

5. Geisslinger, M., Poszler, F., Betz, J., Lütge, C., Lienkamp, M., 2021. Autonomous driving ethics: From Trolley problem to ethics of risk. Philosophy & Technology 1–23. https://doi.org/10.1007/s13347-021-00449-4

6. Gkartzonikas, C., Gkritza, K., 2019. What have we learned? A review of stated preference and choice studies on autonomous vehicles. Transportation Research Part C: Emerging Technologies 98, 323–337. https://doi.org/10.1016/j.trc.2018.12.003

7. Harb, M., Stathopoulos, A., Shiftan, Y., Walker, J.L., 2021. What do we (Not) know about our future with automated vehicles? Transportation Research Part C: Emerging Technologies 123, 102948. https://doi.org/10.1016/j.trc.2020.102948

8. Jo, M., Oh, R., Oh, M.-S., 2023. Prediction of PM10 concentration in Seoul, Korea using Bayesian network. Communications for Statistical Applications and Methods 30, 517–530. https://doi.org/10.29220/CSAM.2023.30.5.517

9. Koski, T., Noble, J., 2011. Bayesian networks: An introduction. John Wiley & Sons. https://doi.org/10.1002/9780470684023

10. Krügel, S., Uhl, M., 2022. Autonomous vehicles and moral judgments under risk. Transportation Research Part A: Policy and Practice 155, 1–10. https://doi.org/10.1016/j.tra.2021.10.016

11. Liu, H., Motoda, H., 2012. Feature selection for knowledge discovery and data mining. Springer Science & Business Media. https://doi.org/10.1007/978-1-4615-5689-3

12. Madden, M.G., 2009. On the classification performance of TAN and general Bayesian networks, in: Bramer, M., Petridis, M., Coenen, F. (Eds.), Research and Development in Intelligent Systems XXV. Springer London, London, pp. 3–16. https://doi.org/10.1007/978-1-84882-171-2_1

13. Martinho, A., Herber, N., Kroesen, M., Chorus, C., 2021. Ethical issues in focus by the autonomous vehicles industry. Transport Reviews 1–22. https://doi.org/10.1080/01441647.2020.1862355

14. Oh, R., Lee, H.K., Pak, Y.K., Oh, M.-S., 2022. An interactive online app for predicting diabetes via machine learning from environment-polluting chemical exposure data. International Journal of Environmental Research and Public Health 19, 5800. https://doi.org/10.3390/ijerph19105800

15. Othman, K., 2021. Public acceptance and perception of autonomous vehicles: A comprehensive review. AI and Ethics 1, 355–387. https://doi.org/10.1007/s43681-021-00041-8

16. Pearl, J., 2009. Causality. Cambridge University Press.

17. Pearl, J. (Ed.), 1988. Probabilistic reasoning in intelligent systems. Morgan Kaufmann, San Francisco (CA). https://doi.org/10.1016/B978-0-08-051489-5.50001-1

18. Radun, I., Radun, J., Kaistinen, J., Olivier, J., Kecklund, G., Theorell, T., 2019. Endangering yourself to save another: A real life ethical dilemma. Transportation Research Part F: Traffic Psychology and Behaviour 64, 318–322. https://doi.org/10.1016/j.trf.2019.05.015

19. Scutari, M., Denis, J.-B., 2021. Bayesian networks: With examples in R, 2nd ed. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9780429347436

20. Sui, T., 2023. Exploring moral algorithm preferences in autonomous vehicle dilemmas: An empirical study. Frontiers in Psychology 14, 1229245. https://doi.org/10.3389/fpsyg.2023.1229245

21. Traffic Lab, 2024. Automated driving system testing and evaluation. Retrieved from: https://traffic.engin.umich.edu/research/automated-vehicle-system-testing-and-evaluation. University of Michigan.

22. Waymo, 2024. Ride with Waymo One. Retrieved from: https://waymo.com/waymo-one/.

23. Xing, Y., Lv, C., Cao, D., Hang, P., 2021. Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving. Transportation Research Part C: Emerging Technologies 128, 103199. https://doi.org/10.1016/j.trc.2021.103199