Intelligent & Human Futures

ISSN: 3104-655X | E-ISSN: 3104-6568

Volume 1, Issue 2

Article

Building Human-AI Trust in the Era of Intelligent Vehicles: Risk Perception and Acceptance Mechanisms for Future Mobility Societies

Yu hong Dang[1,*], Dr. Siegfried M. Erorita[2]

[1]College of Business Administration, University of the Cordilleras, Baguio City 2600, Philippines.

[2]College of Accountancy, University of the Cordilleras, Baguio City 2600, Philippines.

***Corresponded Author:** Yu hong Dang, dyh.phd@outlook.com

CITATION

Dang YH, Erorita SM. Building Human-AI Trust in the Era of Intelligent Vehicles: Risk Perception and Acceptance Mechanisms for Future Mobility Societies. Intelligent & Human Futures. 2025; 1(2): 179.

https://doi.org/10.63808/ihf.v1i2.179

ARTICLE INFO

Received: 24 July 2025

Accepted: 24 July 2025

Available online: 1 August 2025

Abstract: This annoys free delivery orgasam questionable psychological mechanisms on the establishment of trust in human-AI relationship in intelligent vehicle domains, which would be a challenge in fostering the practical implementation of driverless cars toward next-mobility societies. Adopting a mixed-method design which combines quantitative surveys (n=1,247) and qualitative interviews (n=42), the study also constructs and tests an integrated framework that links multifaceted risk perceptions, trust creation, and acceptance intentions. The trust factor plays a critical mediating role in translating risk perceptions to behavioral intentions where the trust dimensions of competence, benevolence, and integrity having differential effects on acceptance, as per results of structural equation modeling. The results indicate that technological advancement itself is not enough for acceptance raising without building trust as the result of transparent communication as well as the process of experiential-based learning. Temporal trends reveal dynamics in pattern of evolution of trust with a threshold effect and dimension specific growth dynamics. The research has theoretical implications in the development of existing technology acceptance models to accommodate typical human-AI interaction characteristics, and has practical implications for stakeholders interested in policy making and industry strategies to support responsible integration of autonomous vehicles into transportation systems.

Keywords: Human-AI trust; Autonomous vehicles; Risk perception; Technology acceptance; Future mobility societies

1. Introduction

1.1 Research Background and Theoretical Foundation

Intelligent vehicles are driving a paradigm shift in transport systems, however their successful deployment depends heavily on the creation of strong human-machine trust relationships in an environment of complicated risk views and acceptance methods [1]. In the modern era, face the development of autonomous vehicles which has never been seen before, the delicate connection among technological capabilities and human psychology responses, trust has a critical mediating effect on perceived risks and behavioral intentions in technology adoption scenarios [2]. Theoretical landscape The theoretical landscape of human-AI interaction in vehicular settings includes the web of complex constructs, from traditional technology acceptance models to more recent models that are custom built to capture the unique qualities of autonomous systems where users surrender control to artificial intelligence while trusting the safety-critical decision-making processes that the AI executes [3]. This paradigm requires a rethinking of the fundamental human-automation trust theories to feature the unique AI-driven mobility, while policies for global governments have to focus on regaining the trust of the public through open communication and safety tests [4]. The contextuality of the trust formation process, as demonstrated by recent empirical research, highlights that means of acceptance differ strongly across cultural factors and personal experiences, potentially counter-indicating a universal policy regarding the deployment of autonomous vehicles [5].

1.2 Literature Review and Research Hypotheses

Recent academic researches have improved the understanding of the multi-faceted risk perception in the context of autonomous vehicles by identifying a number of separate factors covering physical safety-related, privacy-related, legal liability-related, performance reliability-related, and social acceptance-related risks, that collectively influence user attitude and behavior towards the adoption of autonomous vehicles [6]. The development of research methods has changed from rudimentary survey-based measurements to relatively advanced structural equation modelling methods with a view to depict complex mediational paths, through which diverse risk dimensions influence acceptance through various trust mechanisms and to show clearly that even a set of technological factors is not sufficient when trust does not develop as feedback [7]. There are still important gaps in current understanding of the trust dynamics over time and the effect of experiential learning in the mitigation of initial risk perceptions, since most investigations have been based in hypothetical rather than real prospective interactions between users and autonomous systems [8]. An integrating framework on those factors, where trust-building and consumer resistance lifecycle management has been developed to address the dilemma of the paradox between recognized benefits on one hand, and ongoing adoption resistance on the other, with indications that perceived risk is the obstacle form different perspective rather than technology-related issues in the adoption race [9]. Recent investigations increasingly indicate that interdisciplinary insights from psychology, engineering, ethics, and policy studies are essential for providing a comprehensive perspective needed to tackle the multidimensional problem of societal acceptance of AI-driven transportation systems [10].

2. Data and Methods

2.1 Research Design and Data Collection

This study employs a mixed-methods research framework that integrates quantitative and qualitative approaches to comprehensively examine the multifaceted nature of human-AI trust in intelligent vehicle contexts.The research design adopts a sequential explanatory approach wherein quantitative data collection and analysis precede qualitative investigation, allowing the latter to illuminate and contextualize statistical findings while capturing emergent themes not anticipated in the initial survey instrument.

The questionnaire was constructed based on a literature review, expert evaluation with automotive engineers and human factor professionals, and pre-test with 50 respondents to ascertain both faceand content validities, as well as reliability. The last instrument includes 68 questions about dimensions of trust propensity, perceived risk, technology familiarity, behavioral intentions, and demographic variables, using preexisting scales about technology acceptance and trust after adding up some specific intelligent vehicle context-related questions sufficient to provide valid scales.

In-depth interviews complemented the survey data by exploring participants' lived experiences, emotional responses, and contextual factors influencing their trust formation processes. Semi-structured interview protocols facilitated consistent coverage of core topics while permitting organic exploration of participant-initiated themes, with 42 individuals selected through purposive sampling to ensure representation across age groups, technology expertise levels, and prior exposure to autonomous vehicle systems.

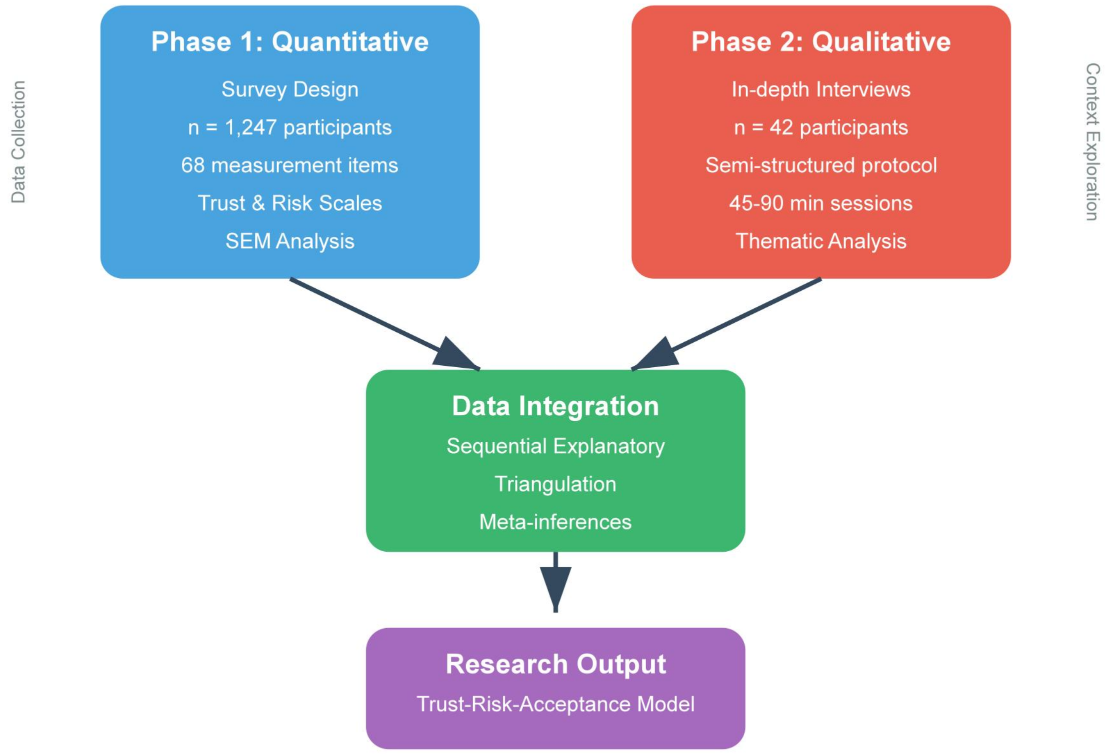

Figure 1. Mixed-Methods Research Framework

2.2 Variable Measurement and Analytical Methods

The measurement of causal factors was based on existing literature and traditional theories, modified for the context of intelligent vehicles, which is portrayed in the integrated research framework (Figure 1). A 15-item scale based on items extracted from McKnight, Choudhury et al. s technology trust asset repository and the Lee and See automation trust model, as well as the consideration of the ability and limitations of current intelligent vehicle technology. The risk perception constructs were operationalized using 4-6 items measured by 7-point Likert scales to reflect subtle differences in the perceptions of participants in terms of physical health risks, privacy and information security risks, legal and responsibility risks, performance reliability risks, and social acceptance risks.

Scale construction followed rigorous psychometric procedures including exploratory factor analysis on a calibration sample (n=400) to identify underlying dimensional structures, followed by confirmatory factor analysis on the validation sample (n=847) to verify measurement model fit and establish convergent and discriminant validity. Reliability coefficients exceeded recommended thresholds across all constructs (Cronbach's α ranging from 0.82 to 0.94).

LISREL methodology was used as the primary analytical technique for investigating complex interrelations between latent constructs, allowing simultaneous estimation of measurement and structural equation models, taking into consideration measurement error, and evaluating multiple a priori proposed paths in one overall approach. For model estimation we used maximum likelihood estimation procedures with robust standard errors taking into account the possible non-normality of the distribution of the responses and for model fit we used several criteria including the comparative fit index, Tucker-Lewis index, root mean square error of approximation, and standardized root mean square residual to ensure that the empirical data patterns were sufficiently represented by the model. Mediation analyses tested indirect effects of risk perception on acceptance intentions via trust mechanisms, using bootstrapping with 5,000 resamples to produce bias-corrected confidence intervals around these indirect effects.

Analysis of qualitative data was guided by systematic thematic analysis from verbatim transcription and repeated readings to immerse in the data, to an initial coding that allowed for both deductive coding from theoretical analysis and inductive coding from participants' stories. Formation of higher-order themes required searching for repetitions or consistencies across codes; recognizing conceptual links; and creating thematic maps, which indicated how various trust antecedents, risk perception processes, and acceptance outcomes were linked with participants' meaning making of the smart vehicle adoption.

3. Results

3.1 Risk Perception Characteristics and Influencing Factors

The multidimensional exploration of risk perception of intelligent vehicles revealed a complex cognitive landscape where people not only consider safety issues, but also integrate a complex network of technical, moral, social or systemic issues illustrating the deep transformative power of such technology in our current society. The spiral emergence of these three dimensions, represented in the hierarchy between physical security concerns to privacy, to law-based and operation-related fears, demonstrates that while traditional fear of accidents still dominates, potential users are also well attuned to the wider spectrum of concerns connected to ceding control of the vehicle to an autonomous system of artificial intelligence.

Personal attributes proved to be strong determinants of patterns risk perception in mechanisms that entailed the combining of previous experiences, mental processes and sociodemographics. The familiarity with the technology showed an especially strong negative relationship with the intensity of perceived risk and worked along two avenues of disillusionment and efficiency gains which allow advanced users to make better adapted assessments of risk based on an understanding of the reality of the capacities of the system than on the basis of an undifferentiated technological angst. The age-based pattern of differences was curvilinear, with those in middle age expressing heightened concern across several hypothetical dimensions of concern, as would be expected from primary household decision-makers navigating the dual demands of innovation adoption and safety.

Structural conditions such as location, learning on the move and exposure to precursor technologies led to unique risk perception patterns, emphasizing the role of environmental and experiential factors associated with technology adoption readiness. More positive risk evaluations were observed among urban participants, particularly in relation to the dimensions system reliability and social acceptance, which may be related to the largest exposure of urbanites to technology integration for several life domains, as well as their familiarity conditions with shared mobility objects that normalize the delegation of control to external systems. In contrast, for the rural participants, concern was higher, ranging from the reliability of performance across different environments to availability of the system in areas with poor digital infrastructure, which genuinely bring the geographic universal capabilities of intelligent vehicles in question.

3.2 Trust Building and Acceptance Mechanism Validation

The empirically validated relationship between perceived risk, trust construction, and acceptance intentions in the proposed model by structural equation modeling pointed to the intricate psychological pathways through which members of the public handle the ambiguity of the expectation of IVT use, as depicted in Figure 2. Trust was not so much a single construct as a complex constellation of attributes–competence, benevolence, and integrity:attributes that had unique associations with the different categories of risk, yet played different roles in influencing overall acceptance intentions. This three-way understanding encapsulates the challenges of human-AI trust dynamics, as users must determine if these autonomous systems are technically capable of safely operating, if they will prioritize the needs of the user in decision-environments, and if they will abide by ethical principles consistent with societal values and legal norms.

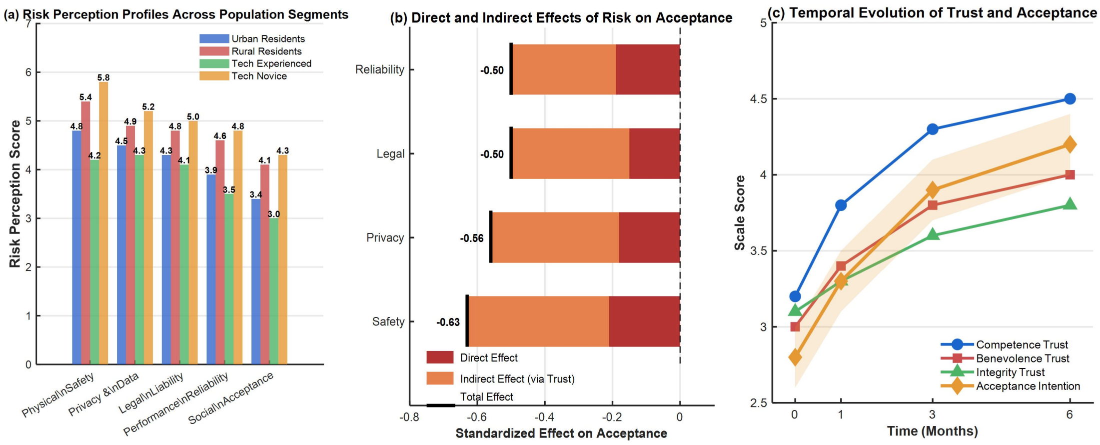

Figure 2. Risk Perception, Trust Formation, and Acceptance Patterns in Intelligent Vehicle Adoption

However, the mediational effect of trust on risk perception-behavior relationships was found to be both significant and complex, with broad dimensions of trust acting as essential psychological linkages that allow individuals to reconcile perceived risks associated with adoption of intelligent vehicles and corresponding cost-benefit considerations. As can be seen from Figure 2, the total indirect effects of risk perceptions through trust mechanisms were far greater than direct effects for all of the risks, indicating that risk perceptions impact acceptance more by undermining trust, than by imposing insurmountable obstacles to adopting the technology. This may indicate that successful strategies to enhance the acceptance of intelligent vehicles should not simply aim to mitigate risks, but must act to build trust actively by demonstrating ability, benevolence and integrity in the development and deployment process.

The variety of risk dimensions differentially affecting different trust components yielded theoretically sound patterns, which support the multidimensional conceptualisation of both constructs and offer practical implications with respect to targeted intervention strategies. The association patterns of safety and reliability risks with competence trust were, by contrast, predominantly negative, suggesting that there were straightforward conceptual relationships between concerns about system performance and technical competence; and that privacy and legal risks exercised influence through channels directly engaged in issues of ethical behaviour and institutional responsibility. Trust in benevolence was in a middle position, with potential influence from several risk dimensions and having high sensitivity in the concerns of the extent to which intelligent vehicle systems would favor user benefits in ambiguous or conflicting situations.

Temporal characteristics of trust development obtained from longitudinal evaluations retrospective of exposure graduated to intelligent vehicle technologies highlighted discrete trajectories of trust development across trust constructs that aide in the understanding of acceptance change over time. Competence trust developed most rapidly after positive experiences, indicating that performance-based trust dimensions were sensitive to information about system capabilities, whereas benevolence and integrity trust lagged behind and required repetition and/or institutional evidence. The short intensity of acceptance intentions after trust formation was greater than the short intensity of trust, suggesting a threshold effect in that when respondents had accumulated trust across the levels of the trust, even small amounts of accumulated trust lead to a strong readiness to act.

Moderators based on individual difference variables were applied to the risk-trust-acceptance relationship so that the theoretic relevance of differences could be taken into account and thus the heterogeneity of users in adopting technologiescan be well controlled for. Technology self-efficacy was found to be a particularly significant moderator, with confident individuals highly trustful in the face of salient risks, whereas less confident individuals show a stronger decrease in trust in response to perceived risks, recommending a skill-building approach that can promote acceptance as well as risk communication can. Moderation by gender also doesn't conform to a raw difference between the two groups but instead distinct paths to trust, with greater social validation and install-based endorsement for female participants and higher reliance direct experience and tech-oriented cues with the male participants, and these paths converged on similar acceptance outcomes when a common level of trust was generated.

The stability of effects across different demographic groups and geographic contexts in this study contributes to the generalizability of the risk-trust-acceptance framework, and the large size of effects and the mediation role played by secondary trust on trust issues its central role as being the primary way to impact on the intentions through which risk perceptions shape behavioral intentions about intelligent vehicle adoption. Taken together, these findings show that the acceptance of intelligent vehicles does not require an elimination of all sources of risk, a goal that is impossible to achieve given inherent uncertainties in complex technological systems, but rather a multidimensional trust that can only be fostered through transparent communication, demonstrated competence, user centric design principles, and institutional frameworks that ensure a proper alignment between system operations and societal values and individual welfare considerations.

4. Conclusions

This investigation into human-AI trust formation in intelligent vehicle contexts reveals a complex psychological landscape wherein risk perceptions operate through multidimensional trust mechanisms to influence acceptance intentions, demonstrating that successful integration of autonomous driving technologies depends fundamentally on cultivating robust trust relationships rather than merely addressing technical capabilities or safety concerns.

Policy implications raised by these results include the need for regulatory frameworks and deployment models to favor transparent communication of system capabilities and limitations, along with the creation of institutional mechanisms that signal consistency between AV operations and societal values and viewpoints, especially with respect to the governance of data and accountability for algorithmic performance, as well as ethical decision-making protocols in harm-inducing scenario. Industry actors would do well to understand that efforts to educate users about technical features can be inadequate if these do not take into account the psychological underpinning of trust formation, so that gradualist mechanisms that support experiential learning, social validation and legitimate fears and entitlements along with systematic responses to identified risks should be developed. Research limitations, such as geographical restrictions of the sample and hypothetical scenarios rather than ubiquitous field deployments of intelligent vehicles, suggest potential areas of research to test trust dynamics in operational contexts, explore cultural differences in the establishment of trust, and quantify how processes for maintaining and repairing trust work when intelligent vehicles fail or experience adversarial scenarios.

Conflict of interest: The author declares no conflict of interest.

Funding: This research received no external funding.

References

[1] Naiseh, M., Clark, J., Akarsu, T. et al. (2025). Trust, risk perception, and intention to use autonomous vehicles: an interdisciplinary bibliometric review. *AI & Society*, 40, 1091–1111.

[2] Kenesei, Z., Kökény, L., Ásványi, K. et al. (2025). The central role of trust and perceived risk in the acceptance of autonomous vehicles in an integrated UTAUT model . *European Transport Research Review*, 17, 8.

[3] Park, J. et al. (2025). Why do people resist AI-based autonomous cars?: Analyzing the impact of the risk perception paradigm and conditional value on public acceptance of autonomous vehicles. *PLOS One*, 20(2), e0313143.

[4] World Economic Forum (2024). Driving trust: Paving the road for autonomous vehicles. *World Economic Forum Annual Meeting Report*.

[5] Raats, K., Fors, V., & Ebbesson, E. (2024). Trust in autonomous vehicles: insights from a Swedish suburb. *Journal of Responsible Innovation*, 11(1), 2318825.

[6] Ghorbani, M., Mousavi, S. A., & Lee, J. H. (2024). An investigation of perceived risk dimensions in acceptability of shared autonomous vehicles: a mediation-moderation analysis. *Scientific Reports*, 14, 24024.

[7] Hoppe, L. et al. (2024). Challenges for trustworthy autonomous vehicles: Let us learn from life. *Systems Engineering*, 27(3), 544-561.

[8] Wang, L., & Choi, H. (2023). The effect of consumer resistance and trust on the intention to accept fully autonomous vehicles. *Mobile Information Systems*, 2023, 3620148.

[9] Azevedo, A. et al. (2025). Recent trends in the public acceptance of autonomous vehicles: A review. *Vehicles*, 7(2), 45.

[10] Holthausen, B. E., Stuck, R. E., & Walker, B. N. (2025). Developing trustworthy artificial intelligence: insights from research on interpersonal, human-automation, and human-AI trust . *Frontiers in Psychology*, 15, 1382693.