Cybersecurity Trust as a Moderator of Perceived Safety Risks in AI-Driven Autonomous Vehicles

Péter Szikora
Keleti Károly Institute of Business Sciences and Digital Knowledge,
Obuda University,
Tavaszmező u. 17, H-1084 Budapest, Hungary
0000-0001-8680-3880

Pál Fehér-Polgár
Keleti Károly Institute of Business Sciences and Digital Knowledge,
Obuda University,
Tavaszmező u. 17, H-1084 Budapest, Hungary
0000-0002-4650-5253

***Abstract*—The social acceptance of self-driving vehicles is closely related to perceived risks and trust in technology, especially in the area of cybersecurity. This study examines whether cybersecurity trust moderates the relationship between perceived risk and acceptance. We conducted an online questionnaire survey in November 2025 (N = 2609, snowball sampling) and used a moderation regression model to analyze the relationships. The results show that perceived risk alone is associated with lower acceptance, while cybersecurity trust is associated with higher acceptance. The moderating effect is significant: with low cybersecurity trust, the negative relationship between risk and acceptance is stronger, while with high trust, this negative relationship weakens significantly, and at the highest levels of trust, a statistically non-negative correlation can be observed. The results suggest that cybersecurity trust is a key factor in the acceptance of autonomous technologies and may weaken the negative association between perceived risks and acceptance.**

Keywords— autonomous vehicles, artificial intelligence, cybersecurity trust, safety perception, technology acceptance

I. INTRODUCTION

The rise of autonomous, AI-powered vehicles is fundamentally changing how society thinks about transportation systems, safety, and the relationship between humans and technology. Alongside the promises of self-driving technologies—such as a reduction in the number of accidents, increased transport efficiency, and more sustainable mobility—the risks and uncertainties that could hinder the widespread adoption of the technology are also becoming increasingly apparent. This is particularly true for systems where decision-making is partly or entirely delegated to artificial intelligence and where the consequences of errors can directly affect human lives. In the case of self-driving vehicles, the concept of safety is not limited to traditional road safety issues. Software reliability, data management practices, and protection against cyber threats have all become integral parts of the system. As a result, information security and safety science are becoming increasingly intertwined, while user confidence plays a key role in how these new types of risks are interpreted and evaluated. Technological safety can therefore not only be measured in objective technical parameters, but also becomes socially relevant through users' subjective risk perception and attitudes of trust.

This study starts from the basic assumption that the acceptance of self-driving vehicles is the result of a complex decision-making process in which perceived safety risks and trust in cybersecurity protection are both decisive factors. The aim of the research is to explore how these factors are associated with the acceptance of autonomous vehicles and whether trust in cybersecurity modifies the negative relationship between perceived risks and acceptance. The study contributes to a deeper understanding of the social acceptance of artificial intelligence-supported systems, with a particular focus on the intersection of safety science and information security. The contribution of the paper is threefold. First, it links the acceptance of autonomous vehicles to both perceived safety risks and cybersecurity-related trust, thereby connecting the literature on technology acceptance with safety science and information security. Second, it empirically tests whether cybersecurity trust moderates the relationship between perceived risk and acceptance, which provides a more nuanced explanation of why risk perception does not always lead directly to rejection. Third, the study uses a large survey sample to identify how subjective risk perception and trust mechanisms jointly shape attitudes toward autonomous vehicles in a contemporary AI-based mobility context.

II. LITERATURE REVIEW

Over the past decade, research into physical and safety risk-taking has increasingly shifted toward multidimensional approaches that go beyond simply examining risk perception. Arslan et al. [1] pointed out that individuals' risk preferences are not directly observable characteristics, but rather constructs based on self-perception, experience, and context. This finding emphasizes the importance of not interpreting risk-taking along a single dimension.

The relationship between risk perception and actual risk-taking has been examined in several empirical studies. Megías-Robles et al. [2] showed that differences between risk-taking and risk perception can be explained in large part by impulsivity and perceived benefits. According to their findings, individuals may be prone to risky behavior even when they accurately recognize the dangers. This approach is reinforced by de Juan-Ripoll et al. [3], who argue that the interpretation of the situation and the individual's propensity for risk-taking together determine domain-specific behavior.

The role of age differences has also become a central topic in risk research. The results of Wang et al. [4] show that younger age groups are more likely to take physical and non-social risks, while risk perception increases and willingness to act decreases

with age. Blankenstein [5] further nuances this picture by pointing out that in certain areas, perceived benefits can offset higher risk perception, leading to the development of heterogeneous behavior patterns.

Profile-based approaches are gaining ground in the empirical study of risk-taking. Fedrigo et al. [6] have shown that external environmental factors can significantly modify risk attitudes and that distinct risk-taking patterns can be identified within the population. These results support the relevance of multivariate approaches that can account for heterogeneous risk-related attitudes.

Technological developments have created new risk contexts that further expand the interpretive framework of risk perception. Lazányi [7] demonstrated in relation to autonomous vehicles that risk perception is not solely linked to technical factors but is also strongly influenced by psychological and social attitudes. This view is reinforced by a recent large-scale study [8], which found that in uncertain environments, risk perception and behavior change adaptively along multiple dimensions.

The relationship between risk perception and safety behavior is the focus of several studies, as confirmed by recent review studies. According to a 2025 systematic review, in jobs involving physical injury, the multidimensional nature of risk perception—including deliberative and affective components—has a decisive influence on safety behavior and individual decisions [9]. At the same time, studies of the relationship between risk and risk perception in adulthood suggest that the effect of age is not evident along a single dimension, but also appears in more subtle patterns of subjective risk perception and behavioral propensity, supporting the use of the propensity models examined [10]. The literature clearly indicates that physical and safety risk-taking is a complex phenomenon that can only be properly understood using an integrated, multivariate approach. A common feature of the empirical and review studies presented is that they treat risk-taking not as a one-dimensional phenomenon, but as a decision-making process in which perceived benefits, intention to act, and risk assessment jointly shape behavior. At the same time, a significant part of the literature focuses on individual dimensions or primarily examines age and contextual differences, while the combined use of integrated indicators and profile-based analysis is less common. This research gap justifies the use of an approach that can simultaneously capture the complex structure of attitudes and the behavioral patterns resulting from individual differences. This theoretical background directly underpins the methodological and analytical framework of the present study.

The above findings provide further motivation for the present study, as the integrated interpretation of risk attitudes is not only a theoretical issue but is directly related to the exploration of different demographic and behavioral profiles. However, previous studies have often examined perceived risk, acceptance, and trust as parallel predictors, while the moderating role of cybersecurity-related trust in the relationship between perceived safety risk and acceptance remains less clearly addressed. The present study contributes to this gap by focusing specifically on whether cybersecurity trust changes the strength and direction of the relationship between perceived risk and acceptance.

III. RESEARCH QUESTIONS AND HYPOTHESES

The social acceptance of self-driving vehicles does not depend solely on their technological sophistication or objective safety indicators, but is also significantly determined by users' subjective risk perceptions and trust in technological systems. The literature increasingly emphasizes that, in the case of safety-critical systems based on artificial intelligence, risk and trust are not independent factors, but influence acceptance attitudes in a complex, mutually reinforcing manner.

In the case of self-driving vehicles, the perceived safety risks are multidimensional, including system failure, loss of control, unexpected decisions, and cybersecurity and data protection threats. These factors are often associated with "black box" technology that is difficult for users to understand, which can increase feelings of uncertainty. Based on this, it can be assumed that higher levels of risk perception are associated with lower acceptance.

At the same time, trust in cybersecurity protection plays a key role in the acceptance of safety-critical systems based on artificial intelligence. In this study, trust is therefore operationalized specifically as trust in cybersecurity protection, not as a comprehensive measure of general trust in autonomous vehicles. If users perceive that self-driving vehicles offer adequate protection against external interference, hacker attacks, and data protection risks, the impact of technological risks on acceptance may be mitigated. This interpretation is consistent with approaches in safety science and information security that focus on human factors, according to which trust can act as a "safety buffer" in the social interpretation of risks.

Based on the above theoretical considerations, the research is based on the following research question and hypotheses:

RQ1: What relationship can be observed between perceived safety risks associated with self-driving vehicles and attitudes toward acceptance, and how does trust in cybersecurity protection influence this relationship?

H1: Higher perceived safety risks related to self-driving vehicles are associated with lower acceptance.

H2: Cybersecurity trust moderates the relationship between perceived safety risks and acceptance, weakening the negative association at higher levels of trust.

Empirical testing of the hypotheses provides an opportunity to explore how trust in information security can shape the social interpretation of safety risks and thus influence the acceptance of self-driving systems. The following chapter presents the methodological framework of the research and the statistical procedures used.

IV. METHODOLOGY

The research data was collected in November 2025 using an online questionnaire. Sampling was not probabilistic, but used a snowball method, in which the initial respondents forwarded the questionnaire to additional respondents through their own networks. This approach is particularly suitable for attitude

surveys related to exploratory, new, and rapidly developing areas of technology—in this case, self-driving vehicles and artificial intelligence-based systems. In particular, snowball sampling may overrepresent respondents from similar social, educational, or technological interest networks, which means that the observed attitudes may partly reflect the characteristics of the recruitment chains rather than the general population.

For the analysis, we used the latest data set collected in November 2025 as part of a larger longitudinal research project, ensuring temporal homogeneity and the examination of attitudes in line with current technological, regulatory, and social discourses. The final sample size was N = 2609, which is considered large in social science attitude research and provides sufficient statistical power for multivariate analyses.

The sample was heterogeneous in terms of gender, age, place of residence, and education, which allowed for a broad mapping of attitudes toward self-driving vehicles. However, due to non-probability sampling, the results cannot be considered representative of the entire population. Therefore, the findings should be interpreted as exploratory and association-based rather than as population-level estimates.

To measure the perceived safety risks associated with self-driving vehicles, we used a five-point scale covering the main risk dimensions related to the operation and safety of the technology, including external interference (hacker attacks), technical failure, unwanted decision-making, risk of loss of control, and concerns about personal data security. Responses were recorded on a five-point Likert scale. The internal reliability of the scale was high (Cronbach's alpha = 0.872), so we used a composite indicator calculated from the average of each item (RISK_MEAN) in our analyses.

The acceptance of self-driving vehicles was measured using three statements that examined purchase intent, willingness to try, and support for the introduction of the technology. Responses were recorded on a five-point Likert scale. The internal reliability of the scale was adequate (Cronbach's alpha = 0.837), and we again created an averaged composite variable (AI_Acceptance) for the analysis.

Trust in cybersecurity protection was included in the model as a moderator variable. This variable measured subjective trust in the information security protection of self-driving vehicles on a four-point ordinal scale. This operationalization captures only one dimension of trust. It does not cover all components of autonomous vehicle trust, such as trust in driving performance, reliability, ethical decision-making, manufacturer responsibility, or institutional oversight. Consequently, the results should be interpreted as referring specifically to cybersecurity-related trust rather than to trust in autonomous vehicles as a whole. In line with empirical social science practice, the four-point cybersecurity trust variable was treated as a quasi-continuous predictor in the analyses. The degree of demand for improvements to the legal and regulatory environment for self-driving vehicles was included as a control variable and was also measured on a four-point Likert scale.

Data analysis was performed using IBM SPSS Statistics 27 software. In the first step, we performed descriptive statistics and reliability tests. We checked the construct validity of the perceived security risk scale using factor analysis, which identified one dominant factor. We used hierarchical moderated linear regression to test the hypotheses. In the first step of the regression model, the perceived risk and the control variable were included, in the second step, cybersecurity trust was included, and in the third step, the interaction term between risk and trust was included. To improve interpretability and reduce multicollinearity, the predictor variables were standardized. To interpret the moderation effect, we performed a simple slopes analysis for low and high levels of trust. We checked the basic assumptions of the regression model with residual diagnostics, which did not indicate any serious violations that would preclude the use of linear regression.

V. RESULTS

The empirical analysis aimed to examine how perceived safety risks associated with self-driving vehicles relate to the acceptance of self-driving technology and what role trust in cybersecurity protection plays in this context. The analyses were limited to data collected in November 2025 (N = 2609) and used a hierarchical moderated regression approach.

Based on descriptive statistics and reliability indices, both the scale measuring perceived safety risk (Cronbach's alpha = 0.872) and the scale measuring acceptance of self-driving vehicles (Cronbach's alpha = 0.837) had adequate internal consistency, allowing composite measures to be used in further analyses.

Table 1 shows the results of the first two steps of the hierarchical regression analysis. In the first model, perceived safety risk and the need for legal regulation were included as predictors. The results show that perceived safety risk was significantly and negatively related to the acceptance of self-driving vehicles ($\beta = -0.190$; $p < 0.001$), while the need for legal regulation showed no significant correlation. The explained variance of the model was low but significant ($R^2 = 0.037$), suggesting that risk perception is relevant in itself but has limited explanatory power.

The analysis strategy was suitable not only for estimating direct associations but also for testing whether cybersecurity trust changes the relationship between perceived safety risk and acceptance. This moderation-oriented approach is consistent with the assumption that acceptance is not shaped by risk perception alone, but by the interaction between perceived risk and protective trust. The descriptive statistics of the main composite variables belonging to the physical/security risk domain are summarized in Table I.

TABLE I. HIERARCHICAL REGRESSION: PREDICTIORS OF ACCEPTANCE OF SELF-DRIVING VEHICLES

Model	Predictor	β	p	R^2
1	Perceived security risk	−0.190	<0.001	0.037
	Need for legal regulation	−0.006	0.759	
2	Perceived security risk	−0.036	0.082	0.145
	Need for legal regulation	−0.017	0.381	
	Cybersecurity trust	0.362	<0.001	

The results show that, on average, respondents characterized the situations examined with low to medium perceived benefits and medium probability of action, while the perceived level of risk was higher. As a result, the average summary risk-taking propensity was moderate.

In the second step, the model was supplemented with the variable of trust in cybersecurity protection. As a result, the explained variance of the model increased significantly (R^2 = 0.145; ΔR^2 = 0.108; $p < 0.001$). Trust in cybersecurity was strongly and positively correlated with acceptance ($\beta = 0.362$; $p < 0.001$), while the effect of perceived risk weakened and did not reach significance in this model. This result suggests that the inclusion of trust substantially modifies the structure of the relationship between risk and acceptance.

In the third step, the model was supplemented with the interaction term between perceived security risk and cybersecurity trust. Table 2 shows the results of the complete moderation model. The interaction effect was significant (β = 0.192; $p < 0.001$), which supports the moderating role of cybersecurity trust in the relationship between perceived risks and acceptance. The explained variance of the total model increased to 18.2% (R^2 = 0.182), which can be considered a medium explanatory power in social science attitude research.

The direction of the interaction is positive, indicating that at low levels of cybersecurity trust, the negative relationship between perceived risk and acceptance is stronger, while at higher levels of trust, this relationship weakens significantly. Based on a simple slope analysis (non-standardized coefficients), in the case of low trust, the increase in risk was associated with significantly lower acceptance ($B = -0.265$; $p < 0.001$), while at high confidence levels, the direction of the relationship changed ($B = 0.145$; $p < 0.001$), indicating a conditionally positive statistical relationship.

TABLE II. MODERATED REGRESSION MODEL FOR THE ACCEPTANCE OF SELF-DRIVING VEHICLES

Predictor	β	p
Perceived security risk	−0.052	0.010
Need for legal regulation	0.001	0.977
Cybersecurity trust	0.364	<0.001
Risk × Trust	0.192	<0.001
R^2	**0.182**	

Based on the empirical results, the hypotheses can be clearly evaluated. The first hypothesis, according to which perceived security risks are negatively related to the acceptance of self-driving vehicles, was conditionally confirmed, as the negative relationship was primarily observed in cases of low cybersecurity trust. The second hypothesis was fully confirmed, as cybersecurity trust played a significant moderating role in the examined relationship.

TABLE III. SUMMARY OF HYPOTHESIS TESTING

Hypothesis	Description	Result
H1	Perceived security risks are negatively associated with the acceptance of self-driving vehicles	Accepted (conditionally)
H2	Cybersecurity trust moderates the relationship between risk and acceptance	Accepted

VI. DISCUSSION

The results of the research are consistent with previous findings that the acceptance of self-driving vehicles is not determined solely by the objective capabilities of the technology, but that users' subjective perceptions of risk and trust in technological systems also play a decisive role. Perceived safety risks are negatively correlated with attitudes toward self-driving technology, which is consistent with well-known risk-acceptance relationships in the field of safety science. According to this approach, higher risk perception is typically associated with lower acceptance, especially in the case of complex and potentially dangerous technological applications.

However, the results also show that trust in cybersecurity protection significantly modifies this relationship. The moderating effect found suggests that trust not only has a direct, positive relationship with acceptance, but can also mitigate the role of perceived risks in reducing acceptance. With low cybersecurity trust, the negative relationship between risk and acceptance is stronger, while at higher levels of trust, this relationship is significantly weakened and a statistically significant relationship in the opposite direction can be observed.

This result offers a possible interpretative framework for the fact that some users do not respond to risk awareness solely with rejection, but with increased interest and a more informed attitude, provided that they perceive adequate cybersecurity protection guarantees. However, it is important to emphasize that this interpretation refers to a hypothetical mechanism that was not directly measured in the present study and therefore requires further empirical investigation.

From a theoretical point of view, the results contribute to the security science interpretation of information security and artificial intelligence-supported systems. In this context, security is not only a technical attribute, but also a social construct in which trust plays a central role. In the case of self-driving vehicles, cybersecurity is not merely background infrastructure, but a key element of perceived system reliability that significantly influences user decision-making. At the same time, cybersecurity trust represents only one component of the broader trust structure surrounding autonomous vehicles. The present findings therefore should not be generalized to all dimensions of trust, but should be understood as evidence for the specific role of perceived cybersecurity protection in shaping acceptance.

The perception of the need for legal regulation did not show a significant correlation with acceptance in the model examined. This result suggests that, in the present sample and within the analytical framework used, respondents' attitudes toward acceptance are more closely related to the direct functioning of the technology and its protective mechanisms than to the formal institutional regulatory environment. This does not mean that regulation is irrelevant, but rather draws attention to the fact that legal frameworks alone are unlikely to be sufficient to increase acceptance if they are not accompanied by perceptible and understandable technical safety solutions.

Overall, the research results suggest that in order to promote the social acceptance of self-driving vehicles, it is not only relevant for developers and decision-makers to reduce technological risks, but also to consciously and transparently build trust in cybersecurity solutions. At the intersection of safety science and information security, trust emerges as a strategic factor that can shape the social interpretation of risks and attitudes toward acceptance.

VII. Conclusion and Implications

The aim of the research was to explore the relationships between perceived safety risks associated with self-driving vehicles, trust in cybersecurity protection, and technological acceptance within the framework of safety science and information security. The empirical results confirm that the acceptance of self-driving vehicles is not solely a function of technological advancement but is significantly shaped by users' risk perception and trust mechanisms.

The results show that perceived artificial intelligence-based security risks are associated with lower acceptance, but this relationship is not deterministic. The moderating role of trust in cybersecurity suggests that at higher levels of trust, the negative association between perceived risks and acceptance is significantly reduced, and a conditionally different statistical relationship can be observed. This result is consistent with the safety science approach, according to which safety is not only reflected in objective technical parameters, but also in subjective social interpretations.

From a practical point of view, the findings suggest that cybersecurity is not merely a background function in the development and introduction of self-driving vehicles, but one of the key factors in user acceptance. Transparent and understandable communication of technical protection solutions can contribute to strengthening trust in the technology and thus indirectly improve attitudes towards acceptance. From a regulatory perspective, the results suggest that legal frameworks alone are not sufficient to increase acceptance if they are not linked to tangible technological safety standards.

The research results must be interpreted with several limitations in mind. Non-probability sampling and the cross-sectional research design limit generalizability and the drawing of causal conclusions. Furthermore, cybersecurity trust was measured with a single item and represented only one dimension of the broader multidimensional trust construct. The study therefore does not capture other forms of trust, such as trust in vehicle reliability, algorithmic decision-making, manufacturers, regulators, or infrastructure providers. Future research may benefit from longitudinal data collection, the use of multi-item trust scales, and the inclusion of experimental or scenario-based approaches. Overall, the study contributes to the safety science and information security interpretation of the social acceptance of autonomous, artificial intelligence-based systems and highlights the strategic role of cybersecurity trust in the development of the acceptance of self-driving vehicles.

Acknowledgment

The authors would like to express their sincere appreciation to the TIG research group for the professional support, constructive collaboration, and continuous guidance provided throughout the research process. The stimulating academic environment and ongoing cooperation substantially contributed to the successful completion of this study.

References

[1] K. Othman, "Public acceptance and perception of autonomous vehicles: A comprehensive review," *AI and Ethics*, vol. 1, no. 3, pp. 355–387, 2021, doi: 10.1007/s43681-021-00041-8.

[2] T. Alqahtani, "Recent trends in the public acceptance of autonomous vehicles: A review," *Vehicles*, vol. 7, no. 2, p. 45, 2025, doi: 10.3390/vehicles7020045.

[3] K. Lazányi, "Perceived risks of autonomous vehicles," *Risks*, vol. 11, no. 2, p. 26, 2023, doi: 10.3390/risks11020026.

[4] F. A. Mir, "An integrated autonomous vehicles acceptance model: Theoretical development and results based on the UTAUT2 model," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 112, pp. 290–304, 2025, doi: 10.1016/j.trf.2025.04.015.

[5] R. de Oña, L. Garach, and J. de Oña, "Social acceptance of autonomous vehicles: A cross-country model validation," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 115, p. 103329, 2025, doi: 10.1016/j.trf.2025.103329.

[6] D. Hafiz, Q. Hou, and I. Zohdy, "Advancing sustainable urban mobility: Public acceptance and perceived risks of autonomous vehicle deployment in Dubai," *Sustainability*, vol. 17, no. 24, p. 11021, 2025, doi: 10.3390/su172411021.

[7] I. Durlik, T. Miller, E. Kostecka, Z. Zwierzewicz, and A. Łobodzińska, "Cybersecurity in autonomous vehicles—are we ready for the challenge?" *Electronics*, vol. 13, no. 13, p. 2654, 2024, doi: 10.3390/electronics13132654.

[8] S. K. Khan, N. Shiwakoti, P. Stasinopoulos, Y. Chen, and M. Warren, "The impact of perceived cyber-risks on automated vehicle acceptance: Insights from a survey of participants from the United States, the United Kingdom, New Zealand, and Australia," *Transport Policy*, vol. 152, pp. 87–101, 2024, doi: 10.1016/j.tranpol.2024.05.002.

[9] R. Szatmáry and K. Lazányi, "Are self-driving cars a safer solution?" in *IFIP International Conference on Human Choice and Computers*, Cham, Switzerland: Springer Nature, 2022, pp. 443–455, doi: 10.1007/978-3-031-47990-8_39.

[10] Z. Kenesei, K. Ásványi, L. Kökény, M. Jászberényi, M. Miskolczi, T. Gyulavári, and J. Syahrivar, "Trust and perceived risk: How different manifestations affect the adoption of autonomous vehicles," *Transportation Research Part A: Policy and Practice*, vol. 164, pp. 379–393, 2022, doi: 10.1016/j.tra.2022.08.022.

[11] Z. Kenesei, L. Kökény, K. Ásványi, and M. Jászberényi, "The central role of trust and perceived risk in the acceptance of autonomous vehicles in an integrated UTAUT model," *European Transport Research Review*, vol. 17, no. 1, p. 8, 2025, doi: 10.1186/s12544-024-00681-x.

[12] M. H. R. Eitrheim, T. Nordfjærn, M. M. Log, and T. Tørset, "Risk perception and trust in autonomous vehicles: A case study of ground handling employees at an international airport," *Journal of Risk Research*, vol. 28, nos. 9–10, pp. 1146–1162, 2024, doi: 10.1080/13669877.2024.2423202.

[13] D. Mohammed and B. Horváth, "Public perception of autonomous vehicles acceptance in Hungary," *International Review of Applied Sciences and Engineering*, vol. 15, no. 3, pp. 435–447, 2024, doi: 10.1556/1848.2024.00769.

[14] F. Nazari, M. Noruzoliaee, and A. K. Mohammadian, "Autonomous vehicle adoption behaviour and safety concern: A study of public perception," *Multimodal Transportation*, p. 100252, 2025, doi: 10.1016/j.multra.2025.100252.

[15] S. Nordhoff, C. Kacperski, P. Liu, and S. Shaheen, "Navigating the road to automated vehicle acceptance," *International Journal of Human–Computer Interaction*, vol. 41, no. 6, pp. 3639–3641, 2025, doi: 10.1080/10447318.2025.2450170.