



Trust in Moral Machines: How automation, morality, and media framing drive cross-cultural adoption of autonomous vehicles

Thomas Teychenié^{a,*}, Julien Cloarec^b, Lars Meyer-Waarden^a

^a *Toulouse School of Management, TSM-Research (CNRS UMR 5303), Université Toulouse Capitole, Toulouse, France*

^b *Iaelyon School of Management, Université Jean Moulin Lyon 3, Magellan, Lyon, France*

ARTICLE INFO

Keywords:

Autonomous vehicles
Automation
Morality
Media framing
Performance trust
Performance risk
Hedonic well-being

ABSTRACT

We conducted a large-scale, tri-national experiment drawing on the Moral Machine clusters—Western individualist (U.S.), Latin-American transitional (Mexico), and East Asian collectivist (China)—to examine how autonomy level (SAE Level 2 vs. Level 5), moral programming (self-protective vs. utilitarian), and accident-severity framing (low vs. high) jointly shape performance trust in autonomous vehicles. In balanced samples of 300 respondents per country, participants were randomly assigned to one of eight conditions and then reported on performance trust, performance risk, hedonic well-being, and behavioral intentions. A robust three-way interaction predicted performance trust in all three contexts, but with opposite patterns: utilitarian framing enhanced the trust advantage of full autonomy under high-severity scenarios in the U.S. and Mexico, whereas in China it did so only when accidents were mild and undermined trust when accidents were severe. These results demonstrate that cultural worldviews condition how ethical programming and risk context interact to shape trust, offering actionable guidance for culturally sensitive AV design and policy.

1. Introduction

Autonomous vehicles are systems capable of performing the entire driving task across all environments and conditions without human input (Mirzakhossein and Mashhadloo, 2024). The degree of independence spans a spectrum defined by the Society of Automotive Engineers, from Level 0, where drivers maintain full control, to Level 5, which requires no human involvement at all (SAE, 2021). Industry leaders such as Tesla, Audi, and Mercedes are racing to refine these systems, driven by the promise of enhanced safety and performance; as a result, the AV market is projected to swell to nearly \$980 billion by 2040 (McKinsey, 2024). Yet consumer acceptance remains fragile. A 2025 survey by the American Automobile Association reveals growing public unease—not only about technical limitations but also about psychological barriers. Foremost among these are ethical anxieties over machines making life-and-death decisions and a pervasive mistrust in the transparency, predictability, and controllability of algorithmic choices (Shariff et al., 2017; Osburg et al., 2022). Trust is further damaged by negative media coverage, especially after AV accidents, which stokes public fears and rapidly erodes trust through the “accident effect” of negative framing (Shariff et al., 2017; Newsweek, 2025; Othman, 2021).

These psychological factors depend not only on external cues such as media narratives, but also on AV design. Crucially, trust in AVs hinges on their moral decision-making in emergencies, where utilitarian programming (sacrificing a few to save many) competes with self-protective programming (prioritizing occupants’ safety) (Bonneton et al., 2016; Bruno et al., 2024). This ethical dilemma—examined in the Moral Machine experiment (Awad et al., 2018)—lies at the heart of consumer hesitation. To counteract these effects, it is essential to foster public trust by increasing transparency around AV ethics—especially the utilitarian versus self-protective trade-off (Bonneton et al., 2016; Bruno et al., 2024)—and by promoting balanced, evidence-based coverage. The MME is invaluable as a descriptive map of moral preferences and cultural patterns. Yet it stops short of explaining the mechanism that connects these preferences to downstream consumer psychology: People’s ethical intuitions about AV decision-making shape their trust, perceptions of risk, sense of well-being behind the wheel, and, ultimately, their intentions to adopt. In other words, while the MME tells us what many people prefer a vehicle to do under moral conflict, it does not specify why those preferences matter for adoption or through which psychological pathways they exert their influence.

Among candidate mechanisms, trust is the most frequently posited

* Corresponding author.

E-mail addresses: thomas.teychenie@tsm-education.fr (T. Teychenié), julien.cloarec@univ-lyon3.fr (J. Cloarec), lars.meyer-waarden@tsm-education.fr (L. Meyer-Waarden).

<https://doi.org/10.1016/j.technovation.2025.103428>

Received 15 November 2024; Received in revised form 6 October 2025; Accepted 15 November 2025

Available online 19 November 2025

0166-4972/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

driver of AV acceptance (Zhang et al., 2019). However, its role is ambivalent. Some studies show that increasing automation enhances trust by improving performance consistency and reducing human error (Dzindolet et al., 2003; Victor et al., 2018; Wu et al., 2023). Other work suggests that higher automation can undermine trust by diminishing users' perceived control, situational awareness, and predictability (Lee and See, 2004; Wang et al., 2020). These mixed findings indicate that trust is not a simple monotonic function of automation. It is shaped by contextual information, including moral expectations about how the vehicle will behave in emergencies (Yokoi and Nakayachi, 2021), and by affective cues such as media framing of accidents that can prime availability heuristics (Kahneman and Tversky, 1979; Shariff et al., 2017).

Yet, trust alone cannot explain those preferences. In trust-based models of autonomous-vehicle adoption, well-being appears to exert a stronger influence on adoption and operates as a downstream consequence of trust (Meyer-Waarden and Cloarec, 2022). Studies confirm that media significantly shape attitudes toward emerging technologies (Liang et al., 2022; Chen et al., 2022), underscoring the importance of guiding these narratives for broader AV adoption. Although AVs hold the potential to revolutionize mobility, psychological barriers continue to impede widespread acceptance, including reduced performance trust, heightened risk perception, and diminished hedonic driver well-being. However, most prior research has centered on classical technology acceptance frameworks (e.g., TAM: Davis, 1989; UTAUT: Venkatesh and Xu, 2012), emphasizing system reliability, usability, and adoption intention. Yet few studies (Yokoi and Nakayachi, 2021) have empirically examined how these moral intuitions influence performance trust, perceived risk, or adoption intentions when users are confronted with AV moral dilemmas framed through media narratives.

Hence, there is a need to assess not only general psychological responses but also the cultural and ethical underpinnings that influence them. Accordingly, we aim to answer the following research questions: How do different levels of AV automation (low vs. high) influence users' performance trust, perceived risk, hedonic well-being and behavioral intention to adopt AVs? How does media framing of AV-related accidents (non-serious vs. fatal crashes) affect users' psychological responses, particularly in terms of performance trust, risk perception, and hedonic well-being? How do performance trust, perceived risk, and hedonic well-being mediate the relationship between AV automation and adoption intentions? How do individual preferences for AV moral programming—utilitarian vs. self-protective—moderate the effects of automation level and media framing on users' performance trust, risks, hedonic well-being, and adoption intention? How do culturally rooted moral intuitions (e.g., from Western, Southern, and Eastern clusters) influence the psychological acceptance of AVs in ethically ambiguous scenarios?

To answer these questions, we conducted an experiment in the three different geographic clusters (Western, Eastern, Southern) defined by the MME, where we manipulated levels of vehicle automation (low vs. high), AV morality (utilitarian vs self-protective) and media-reported accident severity framing (non-serious vs. fatal) through between-subjects experiments ($n = 300$). This study advances cross-cultural AV research by demonstrating that moral cognition, institutional trust, and affective framing jointly condition public responses to autonomous vehicles. Leveraging the Moral Machine's empirically grounded cultural clusters (Awad et al., 2018), we show that Chinese participants—aligned with collectivist, authority-trusting norms—exhibited the highest baseline trust, enjoyment, and adoption intentions, reflecting a techno-optimism shaped by institutional endorsement and communal values. Mexican respondents, consistent with transitional moral logics, expressed moderate trust and heightened perceived risk, suggesting a tension between ethical idealism and infrastructural or institutional uncertainty. U.S. participants, grounded in individualist skepticism and control preferences, showed the lowest trust and weakest behavioral intentions, echoing Western discomfort with moral automation despite theoretical utilitarian endorsement (Bonnenfon et al., 2016; Yoo et al.,

2023). Across all countries, the three-way

Interaction among autonomy, moral logic, and accident severity predicted trust but followed divergent cultural patterns: Western samples favored self-protective logic under low severity, while Chinese respondents favored utilitarian logic—except when faced with high-severity accidents, where trust eroded. These culturally contingent reversals highlight the affective

Volatility of AV trust and align with framing and availability heuristic theories (Kahneman and Tversky, 1979; Shariff et al., 2017), showing how media-amplified accident scenarios override rational appraisals. Finally, trust and risk consistently predicted hedonic well-being and behavioral intentions across samples, affirming their centrality in AV acceptance (Featherman and Pavlou, 2003; Meyer-Waarden and Cloarec, 2022).

We extend the Moral Machine (Awad et al., 2018) literature by exploring how individual moral orientations—specifically preferences for utilitarian or self-protective AV behavior—moderate psychological responses to different automation scenarios. Such parameters have been integrated into studies (Mayer et al., 2021), but to determine what consumers would favor between the two. Takaguchi et al. (2022) examined the two moralities and their impact on consumers' willingness to pay. While Joo and Kim (2023) tied those two moralities to trust and adoption, without considering well-being or different moral clusters. By integrating ethical programming preferences into our model, we shed light on how moral alignment with AV decision-making affects users' trust, well-being, and intentions to adopt. By integrating moral psychology, emotional framing, and cultural cognition, our findings call for localized ethical architectures and communication strategies in global AV deployment.

2. Background

2.1. Autonomous vehicles

Autonomous vehicles are systems capable of performing the entire driving task across all environments and conditions without human input (Mirzahosseini and Mashhadloo, 2024). According to the Society of Automotive Engineers (SAE, 2021), AVs are classified into six levels of automation, each reflecting an increasing degree of autonomy. At Level 0, the vehicle operates like a conventional car, with the human driver responsible for all tasks. Level 1, or driver assistance, introduces basic automation—such as automated steering or acceleration—while keeping the driver in full control. Level 2, partial automation, enables the vehicle to manage both steering and acceleration simultaneously, although the driver must remain attentive and monitor the surroundings. Level 3, conditional automation, allows the vehicle to perform most driving functions under specific conditions, with the driver required to intervene when prompted. Level 4, high automation, enables the car to operate autonomously in most situations, with minimal human oversight needed, typically only in complex or rare scenarios. At Level 5, full automation is achieved: the vehicle can function independently in all environments, without pedals, steering wheels, or any human involvement.

However, as AVs reach full autonomy, they face situations involving life-and-death decisions. Because their actions can directly impact human lives, AVs are increasingly regarded as moral agents (Bonnenfon et al., 2024). Yet the industry is ever pushing for a larger release of those vehicles, as seen with partnership between Waymo and Toyota to give consumers access to it (CNBC, 2025). This technology already in service in some cities in the US is now making its way to Europe (Financial Times, 2025). Indeed, Europe also started testing its own models in Italy (Reuters, 2025). Nevertheless, AV technology raises a critical need to redefine their ethical programming—determining in advance how the vehicle will behave in morally charged scenarios.

2.2. Morality and cultural variability in autonomous vehicle decision-making

A growing body of research investigates the ethical dilemmas associated with AVs, particularly in scenarios where harm is unavoidable and the AVs must choose between conflicting outcomes (Bigman and Gray, 2020). Scholars generally identify two main ethical programming paradigms: a utilitarian approach, which seeks to minimize total harm by prioritizing lives saved, and a self-protective model, which prioritizes the vehicle's occupants (Bonnefon et al., 2016; Bruno et al., 2024). However, public preferences between these paradigms vary widely across individuals and cultures (Bonnefon et al., 2019), raising doubts about the feasibility of implementing a universally accepted moral algorithm. Beyond cross-cultural variation, contextual factors also modulate moral preferences. For example, passengers are more accepting of AVs harming pedestrians than they are of human drivers doing so (Gill, 2020), although this tolerance drops when multiple lives or children are at stake. Mayer et al. (2021) found that while self-protective bias—the tendency to protect those inside the vehicle—remains dominant, utilitarian reasoning becomes more prevalent when higher numbers of lives are involved. However, this recognition does not always translate into action: when asked to configure an AV for themselves or loved ones, most users still prefer self-protective settings (Bonnefon et al., 2016). Even among altruistic individuals, personal proximity and emotional attachment can override utilitarian logic: Yoo et al. (2023) showed that many would choose to protect a family member over ten strangers. Proposals to randomize AV moral decisions to avoid deterministic rules have been largely rejected by consumers, who find such randomness unsettling and unjust (Zhu et al., 2022). Finally, even though AVs remove human agency from the decision loop, many users still feel personally accountable for the machine's actions, especially in high-stakes emergencies. This perceived responsibility reinforces the emotional weight of AV ethical programming, highlighting the need for culturally sensitive, context-aware approaches to AV morality.

This issue was extensively explored in the Moral Machine Experiment (MME), a large-scale global study involving over 490,000 participants from 233 countries (Awad et al., 2018). The MME presented participants with moral dilemmas in which AVs must make critical decisions during unavoidable accidents. These scenarios examined nine moral dimensions (e.g., humans vs. animals, young vs. old, passengers vs. pedestrians, law-abiding vs. jaywalkers) and revealed three distinct moral preference clusters: Western, Southern, and Eastern. Additionally, the MME showed that demographic factors such as age, gender, and level of technological familiarity significantly shaped individual moral choices, highlighting the intricate interplay between cultural norms and personal characteristics in ethical decision-making. Respondents from Western countries (e.g., North America, Western Europe) predominantly endorsed a utilitarian approach, which prioritizes minimizing total harm. They strongly favored saving the greatest number of lives, and showed additional preferences for saving younger individuals over older ones. This reflects a harm-reduction ethic aligned with utilitarianism, where aggregate outcomes—rather than individual roles—drive moral judgments. In Southern countries, primarily in Latin America, participants also favored saving more lives, suggesting a moderate utilitarian orientation. However, their decisions incorporated additional social value judgments, such as prioritizing women and physically fit individuals. This nuanced approach reflects utilitarian thinking tempered by social and possibly cultural notions of vulnerability and utility. By contrast, Eastern countries, which include many Asian and Islamic societies, leaned less toward strict utilitarianism. Instead, their preferences revealed a rule-based or deontological orientation, which emphasized protecting law-abiding individuals and pedestrians over jaywalkers or passengers. This pattern reflects a self-protective and social-order preserving stance, where following rules and maintaining societal norms outweighs maximizing the number of lives saved. While not strictly

"self-protective" in the sense of prioritizing vehicle occupants, this orientation reflects a resistance to utilitarian logic that sacrifices law-abiding individuals in favor of utilitarian calculations.

These findings underscore significant cultural variability in moral reasoning, and illustrate that no single ethical framework—whether utilitarian or self-protective—can be universally applied in the programming of AVs. Moral judgments surrounding AVs are shaped not only by individual and cultural values but also by how ethical dilemmas are framed and contextualized. Genetically encoding AVs with either utilitarian or self-protective logics without acknowledging this complexity risks ethical backlash. A culturally sensitive, context-dependent, and transparently communicated ethical design may be more acceptable to global users than a rigid, universal moral algorithm. These cultural and contextual divergences in moral judgment not only complicate the design of ethical AV algorithms, but also highlight the critical role of how such dilemmas are communicated to the public. Media framing plays a critical role in how AV moral dilemmas are perceived by the public, influencing both emotional reactions and judgments of responsibility. Indeed, 61 % of the US population still fears AVs and this might be due to the intensive accident coverage (eg. BBC, 2023; NBC, 2025; Newsweek, 2025). This brings into focus the importance of media framing, which profoundly shapes public perception and acceptance of autonomous vehicle technologies.

2.3. Media framing and public perception

While moral intuitions surrounding AVs vary across cultural contexts, the way these technologies are communicated to the public plays an equally critical role in shaping perceptions and moral evaluations. Framing theory (Goffman, 1974) posits that individuals rely on interpretive "frames" to make sense of complex information. These frames help structure and contextualize information, shaping how individuals perceive, interpret, and respond to specific issues. Mass media is central in this process, acting as both filter and amplifier. By selecting, emphasizing, or omitting particular aspects of a story, media outlets guide the salience of specific interpretations and steer public opinion (Korneeva et al., 2023). These framing choices often reflect broader cultural values, societal norms, and public sentiment, and contribute to shaping what is seen as legitimate or acceptable within society.

Framing effects have been documented across numerous domains. For instance, media narratives have been shown to influence investor reactions to earnings announcements (Oliver et al., 2023), shape cryptocurrency and fiat currency volatility during crises (Umar et al., 2021), and alter public responses to protests depending on regional context (Li et al., 2025). In particular, emerging and disruptive technologies tend to receive disproportionate negative framing, with media highlighting potential harms more than benefits (Dong et al., 2024). The automotive sector is no exception. For example, the U.S. "Safe Cars Save Lives" campaign demonstrated that increased media coverage could amplify the effectiveness of digital recall campaigns (Pagiavlas et al., 2022). Singh and Grewal (2023) found that while corporate lobbying often weakens recall outcomes, this effect can be offset by high media visibility. In the environmental domain, Chen et al. (2019) showed that media coverage affirming climate change increased hybrid vehicle sales, whereas denialist content had little effect.

In the context of AVs, framing can emphasize either risks—such as technological failures, ethical dilemmas, loss of control and accidents, or benefits, including improved safety, traffic efficiency, and environmental advantages. Such framing significantly could affect public trust, perceived risk, and ultimately, willingness to adopt AV technologies. Dos Santos et al. (2022) identified a trend of consistently negative media coverage, while Zhang et al. (2024) demonstrated that such negativity significantly dampens consumer perceptions and intentions to use AV technologies. Complementing these findings, Cai et al. (2023) showed that sensationalized "over-hype" of AV and ADAS capabilities in media reporting generates widespread safety misconceptions, which—when

unmet by real-world performance—further undermine public trust and willingness to adopt AVs. These findings highlight the potential of media narratives to hinder the public acceptance of AVs, especially when safety concerns dominate headlines.

Despite these insights, key gaps remain. In particular, little is known about how negative media framing of AV-related incidents, especially crashes, influences consumer emotions and behavioral intentions to use AVs. This is especially important given the emotionally charged nature of AV adoption, which implicates personal safety, moral judgment, and well-being. Our research seeks to fill this gap by empirically examining how the framing intensity and emotional tone of media coverage—particularly concerning AV accidents—interact with levels of automation to shape hedonic well-being and adoption intentions. Building on prior sections addressing AV morality and cultural variation, we now turn to our conceptual framework and hypotheses, which model how automation level and media framing jointly affect consumer evaluations and decisions.

3. Conceptual model and hypotheses development

This section develops the conceptual model and hypotheses by synthesizing insights from technology acceptance, trust, risk perception, and well-being research. It integrates moderating influences (e.g., AV morality and media framing) and explores how perceived automation performance shapes cognitive and emotional responses toward AV adoption.

3.1. Effects of AV automation level on performance trust and risk

The degree of automation is a pivotal factor influencing user attitudes toward AVs, particularly in shaping performance trust—defined as the belief that technology will function reliably, credibly, safely, and competently under expected conditions (McKnight and Chervany, 2001; Lee and See, 2004; McKnight et al., 2011). As AVs become more autonomous, users increasingly rely on the system's decision-making capabilities, which can either bolster or undermine trust depending on perceived performance. This trust is especially critical in emerging technologies like artificial intelligence (AI) in AVs, where potential risks and benefits are not fully understood. Recent models of technology acceptance, particularly those focusing on AI (Ostrom et al., 2019; Wirtz et al., 2018; Meyer-Waarden and Cloarec, 2022), underscore the importance of cognitive factors such as performance trust. These factors are essential for mitigating uncertainties inherent in disruptive technologies (Pavlou, 2003), especially when these technologies align with prevention-oriented goals as outlined in regulatory focus theory (Higgins, 1997). In regulatory focus theory, prevention-oriented goals emphasize vigilance, responsibility, and adherence to safety or moral standards—driving individuals to avoid losses and ensure security.

While AVs promise societal benefits—enhanced road safety, reduced emissions, and decreased traffic congestion (Fagnant and Kockelman, 2015; Meyer-Waarden and Cloarec, 2022; Wotton et al., 2022)—real-world evidence supporting the safety benefits of automation alone remains mixed (Noy et al., 2018), creating uncertainty in user perceptions across different automation levels. At lower levels, such as Level 2, drivers must remain actively involved, constantly monitoring the environment and ready to intervene. This high cognitive demand may hinder trust, as users remain acutely aware of the system's limitations (Lee and See, 2004). In contrast, higher levels of automation, such as Level 5, minimize the need for human intervention, leading users to perceive such systems as more capable and autonomous, fostering greater performance trust based on expectations of technical superiority (Dzindolet et al., 2003). However, this confidence can also foster over-trust, where users place excessive reliance on the system despite limited understanding of its true limitations (Victor et al., 2018). Over-trust becomes especially problematic in scenarios where system failure demands immediate human intervention—situations where users may not be

mentally or physically prepared to act due to a false sense of security.

According to the Human-Automation Trust Expectation Model (HATEM) (Carter et al., 2024), user trust is a function of the perceived alignment between the automation's described capabilities and its observed behavior. If Level 5 systems are described as fully autonomous and perform as expected, users are more likely to exhibit strong trust. Conversely, if users encounter underperformance or ambiguous system responses, even high levels of automation may not engender sufficient trust. This is echoed by Rittenberg et al. (2024), who emphasize that trust in automation is strongly influenced by system reliability and user familiarity. Yet even when performance is consistent, higher autonomy often leads users to assume greater competence—amplifying trust regardless of actual reliability (Wu et al., 2023). This perceived competence may stem more from users' mental models than from empirical validation, suggesting that information design and system framing are just as important as functional reliability. Thus, building on the notion that perceived competence is closely linked to automation level, we propose the following hypothesis:

H1a. Higher levels of AV automation positively influence performance trust.

Consumer decisions regarding the adoption of new technologies are often influenced by their perception of risks, which can act as a significant barrier (Ram and Sheth, 1989). According to a study by Statista (2025), the primary concerns surrounding AVs include safety risks due to potential machine errors and the perception that the technology is not yet fully developed. These concerns highlight the importance of cognitive, prevention-oriented factors, such as perceived performance risks, in contemporary models of technology acceptance, particularly in the context of AI and service robots (Ostrom et al., 2019; Wirtz et al., 2018; Meyer-Waarden and Cloarec, 2022; Meyer-Waarden and Cloarec, 2022).

Technology performance risk refers to the extent to which users believe that a technology may increase the likelihood of errors and accidents (Penmetsa et al., 2019; Dogruel et al., 2015; Rosen and Weil, 1994; Sinkovics et al., 2002; Featherman and Pavlou, 2003). It encompasses users' perceptions of their limited ability to effectively manage, control, and safely use an AV, as well as concerns about the AV's reliability in ensuring their safety (Klobas et al., 2019). Technology performance risk presents a significant challenge for AI-driven technologies (Lijarcio et al., 2019). However, there is limited understanding of how technology performance risk affects the adoption of novel and unfamiliar technologies, such as AVs. This risk is often rooted in uncertainty or concerns about the potential consequences of technology failure (Slovic, 1987). Before autonomous driving can be fully realized, it is crucial to gain a comprehensive understanding of the technological risks associated with AV-based interactions (Koopman and Wagner, 2017).

Although AVs are often associated with safety concerns, empirical evidence suggests that users do not inherently perceive them as dangerous. For instance, Nordhoff et al. (2021) found that 80 % of participants reported feeling safe during automated driving, challenging the assumption that automation always increases perceived risk. However, the relationship between automation level and performance risk perception remains under-theorized. As automation advances from Level 2 to Level 5, more driving functions—such as real-time scene interpretation, path planning, and hazard mitigation—are delegated entirely to the AV. At Level 2, users remain actively involved, responsible for monitoring and ready to intervene, which maintains a heightened sense of risk due to their perceived accountability in edge-case scenarios. In contrast, Level 5 systems offer full autonomy, offloading all dynamic decision-making and execution from the driver. This not only reduces cognitive workload but also signals technical competence and reliability, which can lower perceived performance risks (Parasuraman et al., 2000).

This reasoning aligns with Wilde et al. (1985) Risk Homeostasis Theory, which posits that individuals adjust their subjective sense of risk based on perceived safety. When users interact with Level 5 AVs, the

vehicle's end-to-end control over driving tasks leads users to recalibrate their internal risk threshold—they feel safer because the system is perceived as highly capable of managing complex environments. Moreover, as performance trust increases—based on users' belief in the AV's competence—perceived risk tends to decrease. According to the Human-Automation Trust Expectation Model (HATEM) (Carter et al., 2024), greater automation enhances confidence when the system is perceived as sophisticated and dependable. Similarly, Wang et al. (2020) found that reduced perceived risk is contingent on the belief that AVs can autonomously manage unforeseen challenges. Taken together, these findings suggest that users perceive less performance risks at higher automation levels—provided that the system's capabilities align with their expectations of safety and reliability.

H1b. Higher levels of AV automation negatively influence perceived performance risk.

3.2. Moderating role of AV morality: utilitarian vs self-protective approaches

While higher levels of automation generally increase trust in AVs—by reducing human error and enhancing consistency—it remains unclear whether this trust holds when AVs are programmed to follow a utilitarian moral logic. Utilitarianism emphasizes aggregate outcomes, even at the cost of sacrificing individual rights (Bonnefon et al., 2024), and is often perceived as morally unsettling in contexts involving life-or-death decisions.

We define utilitarian orientation as the degree to which an AV prioritizes the greatest good, potentially sacrificing occupants to protect more lives. Research suggests that people judge utilitarian agents—whether human or machine—as less trustworthy, morally rigid, and unpredictable (Everett et al., 2016; Turpin et al., 2021), with these effects consistent across cultures (Everett et al., 2021). In the specific context of AVs, this manifests as skepticism toward moral algorithms that may prioritize pedestrians over passengers in unavoidable crash scenarios (Myers and Everett, 2025).

Although utilitarianism can be associated with moral rationality, in safety-critical and interpersonal contexts like AV decision-making, such logic may conflict with emotional and self-protective human intuitions. Consequently, the trust normally gained from high automation may be undermined when paired with a utilitarian decision framework. Conversely, utilitarianism may heighten perceived performance risk, as passengers fear being sacrificed for a greater good. Indeed, when users learn that an AV follows a utilitarian “sacrifice-one-to-save-many” rule, their personal safety concerns and perceived risk surge, and their willingness to ride declines markedly (Bonnefon et al., 2016). Moreover, even minor deviations from clear moral expectations—such as ambiguous decision rules around whom to protect—can trigger moral aversion that disproportionately inflates perceived risk, above and beyond concerns about technical malfunctions (Dietvorst et al., 2015). Thus.

H2a. Utilitarianism weakens the positive effect of higher levels of AV automation on performance trust.

H2b. Utilitarianism strengthens the negative effect of higher levels of AV automation on performance risk.

3.3. Moderating role of media framing: the case of accident severity

Media framing plays a crucial role in shaping public attitudes toward emerging technologies such as AVs. Framing theory (Kahneman and Tversky, 1979) shows that individuals' interpretations of information are heavily influenced by how that information is presented. Negative news tends to carry more weight than positive coverage (Han et al., 2019), particularly in domains involving risk, safety, or morality. Studies have demonstrated the powerful influence of media in consumer reactions across domains such as food safety (Swartz and Strand, 1981),

hybrid car adoption (Chen et al., 2019), and general technological innovation (Ahluwalia et al., 2000; Borah and Tellis, 2016). In the AV domain, accident reporting—especially when severe—amplifies public concern and erodes trust (Shariff et al., 2017; Othman, 2021). High-profile incidents like the 2016 fatal Tesla crash contributed to heightened caution, akin to public overreactions to rare airline accidents (Graf-Vlachy et al., 2023).

The availability heuristic (Tversky and Kahneman, 1973) explains why vivid, emotionally charged events—such as AV crashes involving fatalities—have disproportionate effects on perceived risk. The more severe the accident, the more salient and memorable it becomes, undermining confidence in AV safety, especially when the vehicle operates at higher automation levels. Even though AVs maintain statistically lower crash rates than human-driven vehicles, highly publicized incidents loom large in people's minds and drive risk perceptions well beyond the data. For instance, in a high-fidelity driving-simulator study, Perello-March et al. (2024) found that participants' self-reported perceived risk and prefrontal cortical activation increased significantly during a sudden hazardous event compared to routine automated driving—demonstrating how vivid, emotionally charged crash simulations amplify risk perceptions. Accident severity may also interact with public perceptions of AV morality. When users are aware that a utilitarian AV could sacrifice them to protect others, reading about a fatal crash may trigger stronger perceptions of vulnerability and distrust. Accordingly, we hypothesize:

H3a. Accident severity conditions the moderated effect of morality on the relationship between AV automation and performance trust.

Meanwhile, perceived risk is likewise shaped by both the probability and magnitude of adverse outcomes. He et al. (2024) introduce a computational model (PCAD) demonstrating that collision severity markedly increases perceived risk under SAE Level 2 automation, and Slovic (1987)'s seminal work on risk perception confirms that individuals' judgments of risk scale with the potential severity of harm. Tversky and Kahneman (1973) explain that the availability heuristic leads consumers to assess event likelihood based on how readily examples come to mind; immediately after reading about a severe AV accident, people may vividly imagine themselves in that scenario. Since utilitarian AVs might sacrifice passengers for the greater good, a fatal crash narrative can cast respondents as unwilling victims of such systems, reinforcing their distrust of AV technology. Thus:

H3b. Accident severity conditions the moderated effect of morality on the relationship between AV automation and perceived performance risk.

3.4. The central role of hedonic well-being

Hedonic well-being refers to the experience of pleasure, comfort, and reduced psychological discomfort (Bauer et al., 2005; Thrash et al., 2010; Waterman, 2008). The use of AVs has the potential to enhance hedonic well-being by improving comfort, reducing travel-related stress by reducing road fatalities through advanced safety features, and offering convenience and enjoyment (Nair and Bhat, 2021; Steg, 2005; Wilson et al., 2022). Performance trust plays a key role in this dynamic (Lee and See, 2004). The relationship between new technologies and hedonic well-being is thus well-established, with research indicating that innovative technologies can contribute to overall well-being by offering new sources of enjoyment and convenience (Pizzi et al., 2019).

Beyond safety and comfort, driving itself carries hedonic appeal where the decision to drive is often driven more by symbolic and emotional factors than by purely practical considerations (Steg, 2005). AVs, with their advanced automation technologies, are perceived not only as practical tools but also as enhancers of personal well-being, providing an effortless and enjoyable driving experience. This perception of increased hedonic well-being is likely to elevate users' intentions

to adopt AV technology (Wilson et al., 2022). Furthermore, technologies that are perceived as hedonic in nature—those that promise pleasure and enjoyment—tend to be more readily embraced by consumers (Pizzi and Scarpi, 2020; Meyer-Waarden and Cloarec, 2022). This is because the anticipation of pleasure plays a critical role in shaping consumer behavior, particularly when adopting new technologies. Thus:

H4. Hedonic well-being positively influences behavioral intentions to use AVs.

Having established how hedonic well-being drives users' intentions to adopt AVs, we now turn to the core cognitive antecedents—performance trust and perceived risk—and examine the factors that shape these pivotal judgments.

3.4.1. Performance trust

According to Lankton et al. (2015), trust comprises three dimensions: system transparency (understanding system logic), technical performance (confidence in reliability), and situation management (ability to regain control). This multidimensional view is essential, as users may cognitively comprehend how an AV works yet still feel anxious or uncomfortable due to perceived lack of control or reliability. Accordingly, in the context of AVs and smart technologies, when users trust the system, they report greater comfort, less anxiety, and a more pleasurable driving experience (Meyer-Waarden and Cloarec, 2022; Jaspers and Pearson, 2022; Meyer-Waarden and Cloarec, 2022). This multi-faceted approach to trust is particularly relevant as it addresses both the cognitive and emotional aspects of user confidence in AV technology, which are not always aligned and can lead to differing outcomes depending on the trust dimension involved.

Research into AI-based AVs (Adnan et al., 2018; Kenesei et al., 2022; Meyer-Waarden and Cloarec, 2022; Zhang et al., 2019) has consistently shown that higher levels of performance trust lead to increased user well-being and a greater likelihood of adopting these technologies. However, this relationship is not merely intuitive. Trust in AVs helps reduce perceived uncertainty and technology-related anxiety—two factors empirically associated with improved hedonic well-being. In the context of AVs, greater trust is theorized to enhance users' hedonic well-being not merely by instilling confidence, but by reducing uncertainty and psychological discomfort—factors empirically shown to improve user experience. Therefore:

H5a. Performance trust positively influences hedonic well-being.

Consequently, this increased trust—through mechanisms such as reduced perceived risk and improved confidence in automation—is also likely to positively influence behavioral intentions to use AVs (Jaspers and Pearson, 2022; Meyer-Waarden and Cloarec, 2022). Thus:

H5b. Performance trust positively influences behavioral intentions to use AVs.

3.4.2. Performance risk

While AVs are engineered to reduce human error and enhance safety, public perceptions often lag behind technical advances. High-profile incidents—such as fatal Uber crashes—have cast doubt on the reliability of fully autonomous systems (Van Brummelen et al., 2018), triggering concerns that such technologies may introduce new types of risk rather than eliminate existing ones. These concerns have made perceived performance risk a central barrier to AV adoption (Wang et al., 2020).

Performance risk refers to users' belief that the AVs may malfunction, behave unpredictably, or expose them to physical or psychological harm (Featherman and Pavlou, 2003; Penmetsa et al., 2019; Dogruel et al., 2015). This includes uncertainty about the vehicle's ability to navigate real-world complexities, detect hazards, and make ethically sound decisions without human intervention (Klobas et al., 2019; Koopman and Wagner, 2017). When users feel that the AV cannot manage such challenges effectively, their well-being is negatively

impacted (Bertrandias et al., 2021).

More broadly, elevated risk perceptions tend to produce technology anxiety, which erodes users' hedonic well-being by increasing discomfort, reducing enjoyment, and fostering a sense of vulnerability (Dogruel et al., 2015; Klobas et al., 2019). In AV contexts, the absence of a human driver may intensify these emotions, especially when users feel uncertain about the system's situational awareness and responsiveness to edge cases (Hohenberger et al., 2016). This erosion of well-being often extends to adoption decisions. Asgari and Jin (2019) showed that perceived risk leads users to resist giving up control—reducing willingness to adopt. Similarly, Featherman et al. (2021) found that performance risk negatively impacts behavioral intentions in the context of electric vehicles, a pattern likely to apply to AVs due to similar concerns over technical maturity and safety.

In parallel, performance risk also undermines performance trust. When users believe an AV may fail under pressure, they are less likely to see it as reliable or capable (Dogruel et al., 2015; Subero-Navarro et al., 2022; Hohenberger et al., 2016). In the case of AVs, safety concerns and incidents, like the Xiaomi, crash intensify distrust by highlighting potential dangers. Additionally, perceived technology performance risk can induce anxiety, further diminishing trust in the technology's reliability or safety. This pattern is observed across various technology domains, including smart home devices and AI systems (Dogruel et al., 2015; Klobas et al., 2019). Trust, as noted earlier, depends on users' perceptions of system competence, reliability, safety, transparency, and controllability (Lankton et al., 2015). As perceived risk increases, these pillars of trust are weakened, resulting in user skepticism and reduced hedonic well-being as well as adoption likelihood (Featherman and Pavlou, 2003; Hohenberger et al., 2016; Klobas et al., 2019; Meyer-Waarden and Cloarec, 2022; Bertrandias et al., 2021). These insights lead to the following hypotheses:

H6a. Performance risk negatively influences performance trust.

H6b. Performance risk negatively influences hedonic well-being.

H6c. Performance risk negatively influences behavioral intentions to use AVs.

Fig. 1 summarizes our research model and the hypotheses.

4. Overview of the studies

We deliberately selected the United States, Mexico, and China because together they capture three distinct moral-cognitive clusters identified in global studies of autonomous-vehicle ethics, while also reflecting divergent market, regulatory, and cultural contexts that shape technology adoption. The Moral Machine project (Awad et al., 2018) clustered societies into Western individualist (e.g., U.S.), Latin-American transitional (e.g., Mexico), and East Asian collectivist (e.g., China) groups based on how people resolve trolley-style moral dilemmas. By including one prototypical country from each cluster, our design ensures that we observe whether deeply held cultural values—such as the emphasis on personal autonomy in the U.S., relational and family obligations in Mexico, and collective welfare in China—condition responses to autonomous-vehicle programming and worst-case accident scenarios.

Beyond moral-dilemma preferences, these three nations also differ markedly in their real-world experiences with, and expectations for, autonomous mobility. The U.S. leads in autonomous-vehicle R&D and regulatory sandboxing, fostering consumer familiarity and trust debates in a mature innovation ecosystem. Mexico represents a burgeoning emerging market where infrastructural challenges, economic constraints, and evolving regulatory frameworks shape pragmatic concerns about safety and accessibility. China, meanwhile, combines rapid deployment of driverless pilot programs with strong state-driven digital ecosystems and collectivist messaging about public welfare. These contrasts allow us to test whether the same ethical programming elicits

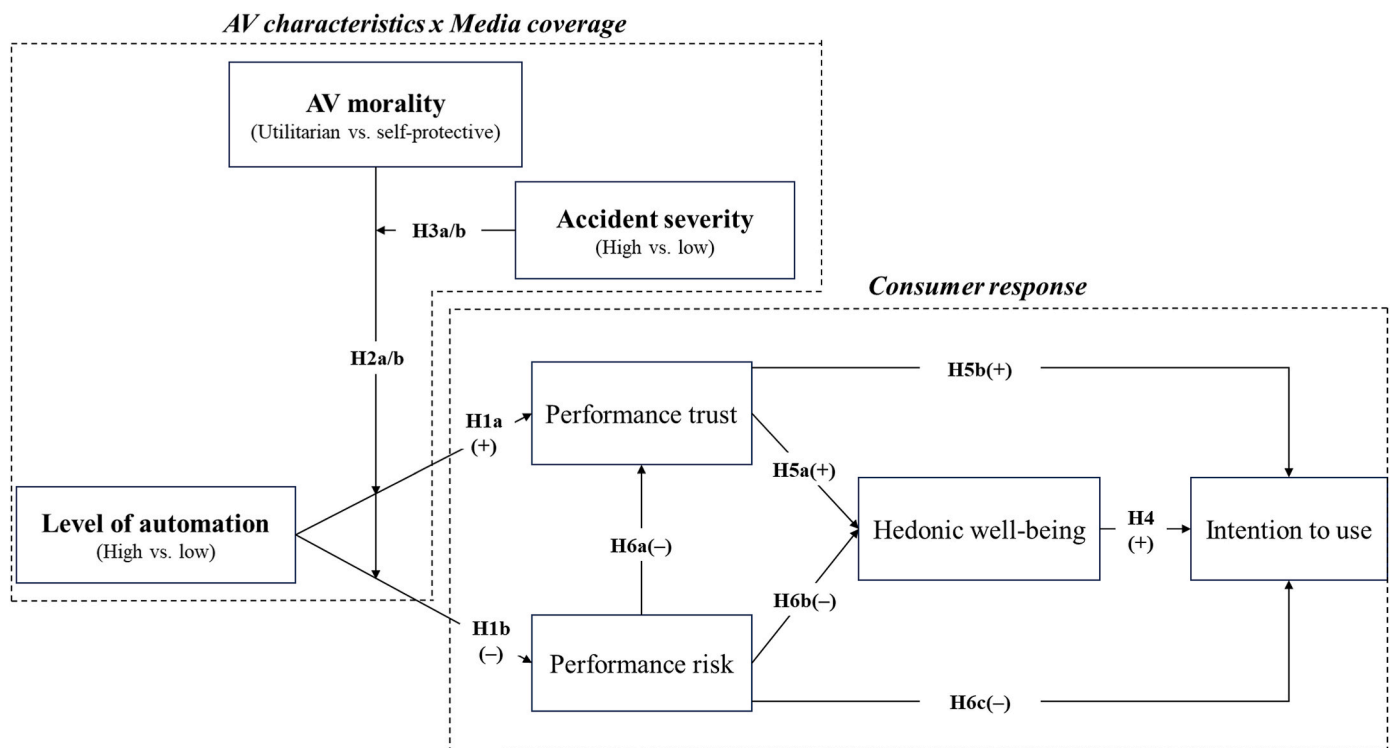


Fig. 1. Research model.

uniform or divergent trust, risk, and adoption intentions across markets at different stages of technological readiness.

Finally, selecting samples of roughly 300 balanced respondents per country (gender-quota matched, language-adapted surveys) not only controls for basic demographics but also maximizes external validity in cross-cultural comparison. By holding age and gender constant while varying only scenario cues, we isolate the extent to which cultural worldviews and lived contexts—rather than demographic noise—drive differences in how people evaluate partial versus full autonomy, self-protective versus utilitarian logic, and low-versus high-severity crash media. This targeted tri-national approach thus yields actionable insights for designing ethically attuned, culturally sensitive autonomous-vehicle systems and for informing policymakers across the globe.

5. Data and methods

5.1. Research design

We implemented a 2 (autonomy level: high vs. low) \times 2 (morality: utilitarian vs. self-protective) \times 2 (accident severity in media coverage: high vs. low) between-subjects experimental design to test our research model. Participants were randomly assigned to one of the conditions, experiencing either a low or high level of vehicle autonomy. In the low-autonomy condition (SAE Level 2), vehicles supported drivers through features such as speed regulation, safe distance maintenance, and steering assistance, but still required driver intervention for critical maneuvers like lane changes or highway exits. These vehicles retained traditional controls (e.g., steering wheel, brake pedals) alongside automated functionalities (SAE, 2021). In contrast, the high-autonomy condition (SAE Level 5) represented full automation, wherein the vehicle managed all driving tasks independently and no longer included manual controls. This level of autonomy, though not yet finalized for public use, is widely anticipated to drastically reduce road fatalities by eliminating human error (Reuters, 2024). The choice of Level 2 and Level 5 AVs reflects both the current reality and projected future of autonomous driving technology. At present, Level 2 systems dominate

the commercial AV market (Statista, 2025), offering driver assistance rather than true autonomy. These systems highlight ongoing user responsibility and allow us to examine how partial control influences user perceptions. Level 5, on the other hand, represents a future paradigm shift, offering a fully autonomous experience independent of human input (SAE, 2021). Comparing these two distinct autonomy levels enables us to capture a broad spectrum of consumer responses to automation, from current technological realities to future expectations (Reuters, 2024).

We also manipulated the moral decision-making logic embedded in the AVs. Participants were told the vehicle either followed a self-protective rule—prioritizing the safety of its occupants—or a utilitarian rule—minimizing total harm, even at the cost of the vehicle's occupants. This moral dichotomy has been widely studied as a central ethical dilemma in AV research (Bonneton et al., 2016; Mayer et al., 2021; Bruno et al., 2024). While the Moral Machine Experiment (Awad et al., 2018) demonstrated the nuance and variability in public moral preferences, a simplified binary framework remains the most tractable approach for examining the broader psychological and behavioral implications of AV moral programming. The self-protective vs. utilitarian contrast provides clear, opposing ethical principles that are conceptually accessible and empirically actionable.

Lastly, we manipulated the severity of a hypothetical AV accident through a media report. Participants read a brief news article describing an AV-related accident that resulted in either no injuries (low severity) or the death of a family with two children (high severity). This decision reflects the real-world media landscape, where AV-related accidents—regardless of outcome—tend to receive heightened and often negative attention (Shariff et al., 2017). While not all AV accidents lead to fatalities, the inclusion of both non-fatal and fatal scenarios allows us to systematically vary the emotional and ethical weight of the incident. Even low-severity incidents may cast AVs in a negative light due to heightened scrutiny, while high-severity incidents underscore the potential for catastrophic outcomes, thereby enabling us to assess how varying degrees of accident severity influence consumer perceptions and trust.

5.2. Samples

We developed three versions of a Qualtrics questionnaire, available in English, Spanish, and Chinese, to accommodate respondents from different cultural and linguistic backgrounds. To ensure diverse and representative samples, we partnered with Prolific, a professional panel provider, to administer the surveys. Specifically, 300 participants from the United States, 300 participants from Mexico, and 300 participants from China completed the respective language versions of the questionnaire. In our recruitment criteria on Prolific, we requested a gender quota, aiming for an approximately equal distribution of male and female respondents within each national sample to enhance demographic representativeness. The characteristics of the samples are summarized as follows: for the United States, 50.7 % of respondents were men, with a mean age of 45.9 years (SD = 15.9); for Mexico, 49.7 % of respondents were men, with a mean age of 30.6 years (SD = 7.4); and for China, 50.7 % of respondents were men, with a mean age of 35.4 years (SD = 10.0). This sampling strategy was intended to facilitate cross-cultural comparisons while controlling for basic demographic factors such as age and gender.

5.3. Measurement instruments

To measure the core constructs of our research model, we used 7-point Likert scales sourced from prior literature. Specifically, performance trust was assessed using items adapted from Kenesei et al. (2022), such as "I can trust that autonomous vehicles can provide a robust and safe mode of transport." To capture performance risk, we also drew on Kenesei et al. (2022), using items such as "Autonomous vehicles may not perform well, and problems may occur when using them." Hedonic well-being was measured using a scale developed by Asante et al. (2024), with items including "I would find riding in autonomous vehicles to be enjoyable," reflecting the emotional and experiential benefits associated with autonomous vehicle use. Finally, behavioral intentions were assessed based on items adapted from Venkatesh and Xu (2012), such as "Looking at its advantages, I intend to buy an autonomous vehicle in the future," to gauge participants' future usage intentions.

5.4. Assessment of the measurement models

The assessment of the measurement model for the United States sample indicates a strong overall fit. The chi-square value ($\chi^2 = 146.879$, $df = 71$) is relatively low, suggesting minimal deviation from the observed data. The Root Mean Square Error of Approximation (RMSEA) is .060, well below the commonly accepted threshold of .08, indicating a good fit. The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) are .987 and .984, respectively, both exceeding the .95 benchmark for excellent model fit. For the Mexican sample, the measurement model also demonstrates an acceptable fit. The chi-square value ($\chi^2 = 197.927$, $df = 71$) is somewhat higher but still within a reasonable range given the sample size and model complexity. The RMSEA is .077, remaining under the .08 threshold and indicating an acceptable fit. The CFI and TLI are .963 and .952, respectively, both surpassing the .95 criterion, supporting a strong model fit. Similarly, the measurement model for the Chinese sample shows a strong fit to the data. The chi-square value ($\chi^2 = 159.181$, $df = 71$) is relatively low, suggesting a good correspondence between the model and the observed data. The RMSEA is .064, well within the acceptable range under .08. The CFI and TLI are .976 and .969, respectively, both exceeding the .95 benchmark, indicating an excellent overall model fit.

The psychometric properties of the measurement scales used in this study were found to be satisfactory. First, the scales demonstrated strong internal consistency reliability, as indicated by Cronbach's alpha values exceeding the commonly accepted threshold of .70 for all constructs (see Table 1). This suggests that the items within each scale reliably measure the intended underlying construct. Convergent validity was also

Table 1
Measurement items, reliability, and convergent validity.

	US		Mexico		China	
	α	AVE	α	AVE	α	AVE
Performance trust (Kenesei et al., 2022)	.95	.87	.77	.55	.86	.69
I can trust that autonomous vehicles can provide a robust and safe mode of transport.						
Driverless cars can be trusted to carry out journeys effectively.						
I trust autonomous vehicles to keep my best interests in mind.						
Performance risk (Kenesei et al., 2022)	.92	.75	.85	.59	.84	.58
Chances are high that something will go wrong when using autonomous vehicles.						
Autonomous vehicles may not perform well, and problems may occur when using them.						
Considering the potential future service performance of autonomous vehicles, their use could be risky for me.						
I'm worried that the failure or malfunctions of autonomous vehicles may cause accidents.						
Hedonic well-being (Asante et al., 2024)	.98	.92	.95	.82	.94	.80
I would find riding in autonomous vehicles to be enjoyable.						
Experiencing autonomous driving technology would be pleasurable.						
I would have fun when traveling in an autonomous vehicle.						
The experience I would gain from using autonomous vehicles would be satisfying.						
Behavioral intentions (Venkatesh and Xu, 2012)	.98	.93	.93	.83	.95	.87
Looking at its advantages, I intend to buy an autonomous vehicle in the future.						
Looking at its benefits, if I had access to an autonomous vehicle, I would intend to buy one.						
The likelihood that I will buy an autonomous vehicle in the future is ...						

established. Specifically, the Average Variance Extracted (AVE) values for each construct were above the recommended minimum criterion of .50 (Table 1), indicating that a substantial proportion of the variance in the observed variables is accounted for by the latent constructs they are intended to measure. Furthermore, discriminant validity was confirmed

Table 2
HTMT coefficients.

		PT	PR	HWB	BI
US	PT	–			
	PR	.78	–		
	HWB	.83	.67	–	
	BI	.83	.67	.82	–
Mexico	PT	–			
	PR	.81	–		
	HWB	.81	.59	–	
	BI	.71	.59	.76	–
China	PT	–			
	PR	.55	–		
	HWB	.82	.45	–	
	BI	.69	.50	.73	–

through the Heterotrait-Monotrait (HTMT) ratio of correlations. All HTMT values fell below the conservative benchmark of .85 (Table 2), supporting the conclusion that the constructs are empirically distinct from one another.

5.5. Common method variance

We confirmed that common method variance was not an issue in this study (Podsakoff et al., 2003). We used the ConMET package (De Schutter, 2021) to test competitive models in which items from two constructs load onto the same latent variable. As shown in Table 3, all configurations significantly reduced the fit of the measurement model, indicated by a significant increase in χ^2 ($p < .001$). Additionally, we evaluated Harman’s one factor (Harman, 1967), which demonstrated inferior performance compared to the measurement model ($p < .001$).

5.6. Post hoc power analysis

Post hoc power analysis is an important procedure for evaluating whether the sample size employed in a study is sufficient to yield robust and reliable estimates (Moshagen and Erdfelder, 2016). Conducting such an analysis helps to ensure that the likelihood of detecting true effects is adequately high, thereby supporting the credibility of the findings. In the present study, we used the semPower package (Jobst et al., 2023) to assess the statistical power of our measurement models. Based on the observed RMSEA values (i.e., .060 for the United States, .077 for Mexico, and .064 for China), as well as a consistent sample size of 300 respondents per country, 71 degrees of freedom for each model, and a significance level (α) set at .05, the results of the power analysis indicated that each model achieved a statistical power (β) of .99. This level of power substantially exceeds the conventional threshold of .80, which is generally considered indicative of sufficient power to detect meaningful effects. Thus, the sample sizes and model specifications used in this study are deemed adequate to support the robustness and reliability of the statistical analyses conducted.

5.7. Method of analysis

To implement the research model and test our hypotheses, we used PROCESS for R, Version 4.0.1, a widely used statistical tool specifically designed for conducting path analyses, as well as moderation and mediation analyses (Hayes, 2022). Developed by Hayes, PROCESS enables researchers to estimate complex models involving direct, indirect, and conditional effects within a flexible and user-friendly framework. Its accessibility and comprehensive functionality have made it a standard tool in the social sciences for testing hypotheses related to mediation, moderation, and conditional process models. In this study, we used a custom syntax to specify our model structure, as detailed below:

Process(data = df, x = "Level", y = "BI", m = c("PR", "PT", "HWB"), w = "Morality", z = "Severity", cov = c("Gender", "Age"), bmatrix = c(1, 1, 1, 0, 1, 1, 0, 1, 1), wzmatrix = c(1, 1, 0, 0, 0, 0, 0, 0, 0), conf = 95).

This syntax outlines the variables included in the model. Specifically, behavioral intentions ("BI") were specified as the dependent variable,

with level of automation ("Level") serving as the independent variable. Morality ("Morality") was included as a moderator, and severity of the accident ("Severity") was modeled as a conditional moderator. The model further incorporated three mediators: performance risk ("PR"), performance trust ("PT"), and hedonic well-being ("HWB"). Additionally, gender and age were included as covariates to account for potential confounding influences. The bmatrix and wzmatrix define the hypothesized relationships among variables and interactions, while the confidence interval for the estimates was set at 95 %.

6. Results

6.1. Descriptive statistics

A one-way ANOVA was conducted to assess whether perceptions of performance trust in autonomous vehicles differed by country (Table 4). There was a significant main effect of country, $F(2, 588) = 37.50, p < .001$. Chinese participants ($M = 4.52, SD = 1.26$) reported significantly greater performance trust than both U.S. ($M = 3.57, SD = 1.73$) and Mexican participants ($M = 3.82, SD = 1.26$), whereas the U.S. and Mexico did not differ from one another.

Performance risk also varied by country: a one-way ANOVA yielded $F(2, 592) = 7.29, p < .001$. Post-hoc Games–Howell comparisons showed that Mexican participants ($M = 5.24, SD = 1.05$) perceived significantly greater performance risk than Chinese participants ($M = 4.92, SD = 1.05$), but neither of these groups differed significantly from the U.S. sample ($M = 5.08, SD = 1.32$).

Hedonic well-being ratings differed across countries as well, $F(2, 586) = 20.92, p < .001$. Chinese respondents again reported the highest sense of hedonic well-being ($M = 4.74, SD = 1.26$), followed by Mexican participants ($M = 4.29, SD = 1.52$) and then U.S. participants ($M = 3.96, SD = 1.79$). All pairwise comparisons were significant (U.S. vs. Mexico, $p < .05$; U.S. vs. China, $p < .001$; Mexico vs. China, $p < .001$).

Finally, behavioral intentions to use autonomous vehicles were compared across the three countries, revealing a significant effect, $F(2, 595) = 44.14, p < .001$. Participants in China expressed stronger intentions ($M = 4.21, SD = 1.66$) than both U.S. ($M = 3.15, SD = 1.88$) and Mexican participants ($M = 3.06, SD = 1.59$), who did not differ significantly from each other.

6.2. Manipulation checks

Manipulation checks confirmed that participants in all three samples perceived the intended differences in autonomy level. In the U.S. sample, a Welch’s *t*-test indicated that those assigned to the high-autonomy (Level 5) condition rated the vehicle as more capable of independent operation ($M = 4.33, SD = 1.99$) than those in the low-autonomy (Level 2) condition ($M = 3.47, SD = 1.76$), $t(295) = 3.95, p < .001$. Mexican participants showed a similarly strong effect under a Student’s *t*-test, $t(298) = 5.02, p < .001$, with means of 4.71 ($SD = 1.44$) versus 3.93 ($SD = 1.24$) for Level 5 and Level 2 respectively. In China, a Welch’s *t*-test again confirmed the manipulation—Level 5 participants perceived greater autonomy ($M = 4.30, SD = 1.73$) than Level 2 participants ($M =$

Table 3
Common method variance estimation.

		US		Mexico		China	
	df	χ^2	$\Delta\chi^2$	χ^2	$\Delta\chi^2$	χ^2	$\Delta\chi^2$
Proposed Model	71	146.879		197.927		159.181	
PT and PR	74	511.48	364.601***	285.118	87.191***	488.064	328.884***
PT and HWB	74	644.477	497.598***	339.628	141.702***	326.845	167.664***
PT and BI	74	638.937	492.057***	402.361	204.434***	509.086	349.906***
PR and HWB	74	752.844	605.965***	536.776	338.849***	565.199	406.018***
PR and BI	74	739.244	592.364***	528.733	330.806***	540.38	381.199***
HWB and BI	74	888.104	741.225***	576.276	378.349***	751.329	592.149***
Harman’s Factor	77	1760.364	1613.484***	925.919	727.993***	1246.217	1087.036***

Table 4
Descriptive statistics results.

	US		Mexico		China		Welch's test	Games-Howell's test		
	M	SD	M	SD	M	SD		US vs Mexico	US vs China	Mexico vs China
PT	3.57	1.73	3.82	1.26	4.52	1.26	F(2, 588) = 37.50***	ns	***	***
PR	5.08	1.32	5.24	1.05	4.92	1.05	F(2, 592) = 7.29***	ns	ns	***
HWB	3.96	1.79	4.29	1.52	4.74	1.26	F(2, 586) = 20.92***	*	***	***
BI	3.15	1.88	3.06	1.59	4.21	1.66	F(2, 595) = 44.14***	ns	***	***

3.74, SD = 1.52), t(292) = 2.98, p = .003.

Checks of the morality framing also demonstrated clear differentiation across conditions. In the U.S., a Student's t-test revealed that participants in the utilitarian framing rated the decision rule as more focused on minimizing overall casualties (M = 5.24, SD = 1.73) than those in the self-protective framing (M = 3.61, SD = 1.82), t(298) = 7.95, p < .001. Similarly, Mexican participants distinguished utilitarian (M = 5.09, SD = 1.79) from self-protective logic (M = 3.52, SD = 1.75), t(298) = 7.69, p < .001. In China, the effect was smaller but still significant, t(298) = 2.84, p = .005, with means of 4.44 (SD = 2.08) versus 3.76 (SD = 2.05) for utilitarian and self-protective framings respectively.

Finally, the accident-severity manipulation was effective in all three contexts. U.S. participants exposed to the high-severity report rated the accident as more severe (M = 5.51, SD = 1.91) than those who read the low-severity report (M = 3.26, SD = 1.73), t(298) = 10.60, p < .001. In Mexico, a Welch's t-test yielded t(222) = 10.60, p < .001, with means of 6.56 (SD = .77) for high severity and 3.65 (SD = 1.60) for low severity. Chinese participants likewise perceived the high-severity scenario as significantly more severe (M = 6.14, SD = 1.31) than the low-severity scenario (M = 4.28, SD = 1.56), t(282) = 11.10, p < .001. Thus, all three manipulations operated as intended across countries.

6.3. ANCOVAs

6.3.1. Performance trust

Across all three countries, the three-way interaction (Level × Morality × Severity) significantly predicted performance trust—though in opposite directions in China versus the U.S. and Mexico. In the U.S. and Mexico, utilitarian framing amplified the trust advantage of Level 5 only under high-severity scenarios; in China, utilitarian logic boosted trust in low-severity contexts but eroded it when severity was high.

In the U.S. sample, a 2 (autonomy level) × 2 (morality) × 2 (severity) ANCOVA on performance trust revealed that the main effect of autonomy level was not significant, F(1, 289) = .83, p = .36, indicating no support for H1. The interaction between autonomy level and morality framing also failed to reach significance, F(1, 289) = 1.11, p = .29, providing no support for H2. However, the three-way interaction among autonomy level, morality framing, and accident severity was significant, F(1, 289) = 7.26, p = .007, supporting H3 and suggesting that the combined influence of moral logic and accident severity on performance trust depends on the vehicle's autonomy level.

In the Mexican sample, results mirrored those in the U.S. First, the main effect of autonomy level on performance trust was non-significant, F(1, 289) = .67, p = .31, again contradicting H1. The autonomy × morality interaction was likewise non-significant, F(1, 289) = .07, p = .79, offering no support for H2. The three-way interaction was highly significant, F(1, 289) = 13.86, p < .001, consistent with H3 and indicating that the interplay of autonomy, moral framing, and accident severity shapes trust in a way that cannot be captured by any two-factor combination alone.

In the Chinese sample, the main effect of autonomy level approached, but did not attain, significance, F(1, 289) = 3.71, p = .06, so H1 was not supported. The autonomy × morality interaction remained non-significant, F(1, 289) = .82, p = .37, offering no support for H2. Crucially, the three-way interaction was significant, F(1, 289) =

5.94, p = .02, again supporting H3 by demonstrating that the combined effects of autonomy level, moral decision rule, and accident severity on performance trust are interdependent.

6.3.2. Performance risk

Across all three country samples, none of the hypothesized effects (main or interaction) of autonomy level, moral framing, or accident severity on performance risk were supported.

In the U.S. sample, a 2 (autonomy level) × 2 (morality) × 2 (severity) ANCOVA on performance risk revealed that the main effect of autonomy level was not significant, F(1, 290) = 1.30, p = .26, indicating no support for H1. The interaction between autonomy level and morality framing also failed to reach significance, F(1, 290) = 2.92, p = .09, providing no support for H2. Finally, the three-way interaction among autonomy level, morality framing, and accident severity was non-significant, F(1, 290) = .02, p = .90, indicating no support for H3 in the U.S. context.

In the Mexican sample, the main effect of autonomy level on performance risk was likewise non-significant, F(1, 290) = .72, p = .40, contradicting H1. The autonomy × morality interaction yielded F(1, 290) = .35, p = .56, again offering no support for H2. The three-way interaction was effectively null, F(1, 290) = .00, p = .98, providing no evidence in favor of H3 in the Mexico sample.

In the Chinese sample, autonomy level had no significant effect on perceived performance risk, F(1, 290) = 1.35, p = .25, so H1 was not supported. The interaction between autonomy level and morality framing was also non-significant, F(1, 290) = .00, p = .98, providing no support for H2. The three-way interaction among autonomy level, morality framing, and accident severity likewise failed to reach significance, F(1, 290) = .02, p = .89, yielding no support for H3 in the China context.

6.4. Model estimation

The model was tested separately in each country sample using PROCESS (Table 5). Across all three samples, the direct effect of autonomy level on performance trust (H1a) was positive but reached significance only in Mexico (b = .68, p < .001), whereas it was non-significant in the U.S. (b = .32, p = .12) and China (b = .27, p = .10). The effect of autonomy level on performance risk (H1b) was uniformly non-significant in all three contexts (U.S. b = -.15, p = .08; Mexico b = .08, p = .40; China b = .02, p = .82).

The two-way interaction between level and morality on performance trust (H2a) did not reach significance in the U.S. (b = -.43, p = .18) or China (b = .39, p = .09), but was significant in Mexico (b = -.81, p < .01), indicating that the impact of autonomy level on trust in Mexico depended on the ethical framing of the vehicle's decision rule. In contrast, the level × morality interaction on performance risk (H2b) was non-significant across all samples (all ps > .10).

The three-way interaction among level, morality, and severity for predicting performance trust (H3a) was significant in all three countries, though with differing signs in China: U.S. b = 1.42, p < .01; Mexico b = 1.50, p < .001; China b = -1.25, p < .05. This suggests that the combined effects of autonomy level, moral logic, and accident severity jointly shape trust, but in opposite directions in China versus the Western samples. The analogous three-way interaction on performance

Table 5
Model estimation results.

Research model	Effect		
	US	Mexico	China
Hypotheses			
H1a. Level → PT	.32ns	.68***	.27ns
H1b. Level → PR	-.15ns	.08ns	.02ns
H2a. Level x Morality → PT	-.43ns	-.81**	.39ns
H2b. Level x Morality → PR	.46ns	.13ns	-.03ns
H3a. Level x Morality x Severity → PT	1.42**	1.50***	-1.25*
H3b. Level x Morality x Severity → PR	.07ns	.01ns	.07ns
H4. HWB → BI	.42***	.56***	.62***
H5a. PT → HWB	.73***	.69***	.70***
H5b. PT → BI	.44***	.16*	.26**
H6a. PR → PT	-.90***	-.71***	-.56***
H6b. PR → HWB	-.15*	-.22**	-.10*
H6c. PR → BI	-.11ns	-.25**	-.26***
Control variables			
Gender → PT	.01ns	.19ns	.05ns
Gender → PR	-.50***	-.13ns	-.13ns
Gender → HWB	.13ns	.31*	-.11ns
Gender → BI	.01ns	.04ns	.10ns
Age → PT	-.02***	-.00ns	.00ns
Age → PR	.01ns	.00ns	-.02*
Age → HWB	-.01*	-.01ns	-.00ns
Age → BI	-.01**	.00ns	.00ns

risk (H3b) was non-significant in every context (all ps > .50).

Turning to downstream outcomes, hedonic well-being (HWB) robustly predicted behavioral intentions in all three samples (H4: U.S. b = .42, p < .001; Mexico b = .56, p < .001; China b = .62, p < .001). Performance trust exerted a strong positive effect on hedonic well-being in every country (H5a: U.S. b = .73, p < .001; Mexico b = .69, p < .001; China b = .70, p < .001), and also directly influenced behavioral intentions (H5b), with significant paths in the U.S. (b = .44, p < .001), Mexico (b = .16, p < .05), and China (b = .26, p < .01).

Finally, performance risk negatively influenced performance trust across all samples (H6a: U.S. b = -.90, p < .001; Mexico b = -.71, p < .001; China b = -.56, p < .001). Performance risk also had a small but significant negative effect on hedonic well-being (H6b) in each country (U.S. b = -.15, p < .05; Mexico b = -.22, p < .01; China b = -.10, p < .05). Its direct impact on behavioral intentions (H6c) was non-significant in the U.S. (b = -.11, p = .12) but significant and negative in Mexico (b = -.25, p < .01) and China (b = -.26, p < .001).

To further probe the significant three-way interaction on performance trust, we examined the simple autonomy-by-morality interaction at low and high levels of accident severity in each country (Table 6). In the U.S. sample, under low-severity conditions the autonomy × morality interaction on performance trust was not significant (b = -.43, F(1, 289) = 1.33, p > .05), indicating that when the accident scenario was mild, framing the vehicle’s logic as utilitarian versus self-protective did not alter the (nonexistent) trust advantage of high autonomy. However, when severity was high, the interaction became positive and significant (b = .99, F(1, 289) = 6.87, p < .01), such that utilitarian framing amplified the trust benefit of a Level 5 vehicle relative to a Level 2 vehicle in the context of severe accidents.

In Mexico, a different pattern emerged: at low severity the autonomy

Table 6
Test of conditional interaction (Level x Morality) at value of severity.

Country	Severity	Effect	F
US	Low	-.43ns	F(1, 289) = 1.33ns
	High	.99**	F(1, 289) = 6.87**
Mexico	Low	-.81**	F(1, 289) = 8.10**
	High	.70*	F(1, 289) = 5.79*
China	Low	.39ns	F(1, 289) = 1.13ns
	High	-.86*	F(1, 289) = 5.82*

× morality interaction was significantly negative (b = -.81, F(1, 289) = 8.10, p < .01), suggesting that utilitarian logic actually reduced the trust advantage of the high-autonomy vehicle when the accident report was mild. Yet under high-severity conditions, the interaction reversed sign and reached significance in the positive direction (b = .70, F(1, 289) = 5.79, p < .05), indicating that utilitarian framing bolstered the trust associated with Level 5 autonomy when the accident was depicted as severe.

In the Chinese sample, moral framing did not significantly modulate the trust effect of autonomy at low severity (b = .39, F(1, 289) = 1.13, p > .05). However, at high severity the interaction became significantly negative (b = -.86, F(1, 289) = 5.82, p < .05), indicating that when reading about a severe accident, utilitarian framing actually weakened the performance trust advantage of the fully autonomous vehicle relative to the self-protective logic.

In the U.S. sample, the three-way interaction among autonomy level, morality framing, and accident severity on performance trust was significant, F(1, 289) = 7.26, p < .01 (Table 7). Simple-slopes analysis revealed that under self-protective framing, autonomy level had no appreciable effect on trust when severity was low (b = .32, ns), but exerted a significant negative effect when severity was high (b = -.84, p < .01), indicating that in severe-accident scenarios a self-protective logic actually diminished the trust advantage of higher autonomy. Under utilitarian framing, autonomy level did not significantly influence trust at either low (b = -.11, ns) or high severity (b = .15, ns), suggesting that utilitarian logic neutralized autonomy’s impact on performance trust regardless of accident severity.

In the Mexican sample, the Level × Morality × Severity interaction was highly significant, F(1, 289) = 13.86, p < .001. When the vehicle was framed as self-protective, higher autonomy significantly increased trust in low-severity contexts (b = .68, p < .001), but actually reduced trust in high-severity contexts (b = -.46, p < .05). Conversely, under utilitarian framing, autonomy level had no significant effect on trust at low severity (b = -.12, ns) or high severity (b = .23, ns), indicating that utilitarian logic again buffered performance trust against changes in autonomy when accident outcomes varied.

In China, the three-way interaction also reached significance, F(1, 289) = 5.94, p < .05. For self-protective framing, autonomy level did not significantly affect trust at low (b = .27, ns) or high severity (b = .46, ns). Under utilitarian framing, however, autonomy level positively influenced trust in low-severity scenarios (b = .66, p < .05) but not in high severity (b = -.40, ns). Thus, whereas Western samples showed a reversal of the autonomy-trust relationship under severe accidents only for self-protective logic, the Chinese sample exhibited a trust-boosting effect of higher autonomy under utilitarian logic—but only when the accident was mild.

6.5. Conditional moderated mediation analysis

In the U.S. sample (Table 8), the conditional indirect effect of

Table 7
Conditional effects of level at values of the moderators.

Country	Level*Morality*Severity	Morality	Severity	Effect
US	F(1, 289) = 7.26**	Self-protective	Low	.32ns
		Self-protective	High	-.84**
		Utilitarian	Low	-.11ns
		Utilitarian	High	.15ns
Mexico	F(1, 289) = 13.86***	Self-protective	Low	.68***
		Self-protective	High	-.46*
		Utilitarian	Low	-.12ns
		Utilitarian	High	.23ns
China	F(1, 289) = 5.94*	Self-protective	Low	.27ns
		Self-protective	High	.46ns
		Utilitarian	Low	.66*
		Utilitarian	High	-.40ns

Table 8
Conditional moderated mediation results.

Country	Path	Morality	Severity	Effect	95 % CI		Index	
					Lower	Upper		
US	PT	Self-protective	Low	.14ns	-.0719	.3643	.62 [1.777, 1.1614]	
		Self-protective	High	-.37*	-.6397	-.1291		
		Utilitarian	Low	-.05ns	-.3452	.2137		
		Utilitarian	High	.07ns	-.1286	.2702		
	PT + HWB	Self-protective	Low	.10ns	-.0509	.2538		.44 [.1158, .8275]
		Self-protective	High	-.26*	-.4695	-.0825		
		Utilitarian	Low	-.03ns	-.2387	.1450		
		Utilitarian	High	.05ns	-.0865	.1857		
Mexico	PT	Self-protective	Low	.11*	.0017	.2578	.24 [.0048, .5456]	
		Self-protective	High	-.07ns	-.2093	.0064		
		Utilitarian	Low	-.02ns	-.1003	.0429		
		Utilitarian	High	.04ns	-.0238	.1154		
	PT + HWB	Self-protective	Low	.27*	.1120	.4351		.59 [.2715, .9559]
		Self-protective	High	-.18*	-.3827	-.0034		
		Utilitarian	Low	-.05ns	-.2117	.0967		
		Utilitarian	High	.09ns	-.0584	.2402		
China	PT	Self-protective	Low	.07ns	-.0730	.2099	-.33 [-.7756, -.0433]	
		Self-protective	High	.12*	.0016	.2897		
		Utilitarian	Low	.17*	.0297	.3952		
		Utilitarian	High	-.10ns	-.3106	.0469		
	PT + HWB	Self-protective	Low	.12ns	-.0971	.3430		-.55 [-1.0358, -.1110]
		Self-protective	High	.20*	.0057	.4357		
		Utilitarian	Low	.29*	.0799	.5271		
		Utilitarian	High	-.17ns	-.4520	.0735		

autonomy level on behavioral intentions via performance trust alone was positive but non-significant under self-protective framing when severity was low ($b = .14$, 95 % CI = $[-.0719, .3643]$). When severity was high, this indirect effect became significantly negative ($b = -.37$, 95 % CI = $[-.6397, -.1291]$, $p < .05$), indicating that under severe accident conditions and a self-protective logic, higher autonomy actually reduced intentions via lowered trust. Under utilitarian framing, neither low ($b = -.05$, 95 % CI = $[-.3452, .2137]$) nor high severity ($b = .07$, 95 % CI = $[-.1286, .2702]$) yielded significant indirect effects. The index of moderated mediation for the PT path was .62 (95 % CI = $[.1777, 1.1614]$), confirming that the severity-by-morality interaction significantly altered the indirect trust-based pathway to intentions.

When hedonic well-being was added as a second mediator, the pattern was similar but attenuated. Under self-protective framing with low severity, the sequential indirect effect through PT then HWB was non-significant ($b = .10$, 95 % CI = $[-.0509, .2538]$). Under self-protective/high-severity conditions, the indirect effect was significantly negative ($b = -.26$, 95 % CI = $[-.4695, -.0825]$, $p < .05$). Utilitarian framing again produced no significant effects at low ($b = -.03$, 95 % CI = $[-.2387, .1450]$) or high severity ($b = .05$, 95 % CI = $[-.0865, .1857]$). The index of moderated mediation for the PT→HWB path was .44 (95 % CI = $[.1158, .8275]$), indicating that the conditional indirect effect through both mediators varied significantly by morality and severity.

In Mexico, the indirect effect via performance trust was small but significant under self-protective framing when severity was low ($b = .11$, 95 % CI = $[.0017, .2578]$, $p < .05$), suggesting that higher autonomy increased intentions by boosting trust in mild-accident/self-protective contexts. This effect reversed to non-significant under self-protective/high-severity conditions ($b = -.07$, 95 % CI = $[-.2093, .0064]$), and remained non-significant under utilitarian framing for both low ($b = -.02$, 95 % CI = $[-.1003, .0429]$) and high severity ($b = .04$, 95 % CI = $[-.0238, .1154]$). The index of moderated mediation was .24 (95 % CI = $[.0048, .5456]$).

For the sequential PT→HWB pathway, self-protective/low-severity conditions yielded a significant positive indirect effect ($b = .27$, 95 % CI = $[.1120, .4351]$, $p < .05$), whereas self-protective/high-severity conditions produced a significant negative effect ($b = -.18$, 95 % CI = $[-.3827, -.0034]$, $p < .05$). Under utilitarian framing, neither low ($b = -.05$, 95 % CI = $[-.2117, .0967]$) nor high severity ($b = .09$, 95 % CI = $[-.0584, .2402]$) showed significant indirect effects. The moderated mediation index for the two-mediator path was .59 (95 % CI = $[.2715, .9559]$).

In the Chinese sample, the indirect effect via performance trust was non-significant under self-protective/low-severity scenarios ($b = .07$, 95 % CI = $[-.0730, .2099]$), but became significant and positive under self-protective/high-severity conditions ($b = .12$, 95 % CI = $[.0016, .2897]$, $p < .05$). Under utilitarian framing, the indirect effect was significant and positive at low severity ($b = .17$, 95 % CI = $[.0297, .3952]$, $p < .05$) but non-significant at high severity ($b = -.10$, 95 % CI = $[-.3106, .0469]$). The index of moderated mediation for the PT path was $-.33$ (95 % CI = $[-.7756, -.0433]$), indicating a significant conditional indirect effect in the opposite direction compared to Western samples.

When including HWB as a second mediator, none of the self-protective/low-severity ($b = .12$, 95 % CI = $[-.0971, .3430]$) or utilitarian/high-severity ($b = -.17$, 95 % CI = $[-.4520, .0735]$) paths reached significance, but both self-protective/high-severity ($b = .20$, 95 % CI = $[.0057, .4357]$, $p < .05$) and utilitarian/low-severity ($b = .29$, 95 % CI = $[.0799, .5271]$, $p < .05$) conditions produced significant positive indirect effects. The index of moderated mediation for the PT→HWB path was $-.55$ (95 % CI = $[-1.0358, -.1110]$), confirming that conditional indirect effects through both mediators varied significantly by morality framing and accident severity in China.

7. Discussion

7.1. Theoretical implications

These cross-national patterns underscore how deeply rooted cultural logics shape baseline attitudes toward AVs even before ethical framing or accident scenarios come into play. In China—a prototypical collectivist context where technological stewardship often enjoys broad institutional endorsement—participants entered the study with notably higher confidence in automated driving systems, felt they would derive more enjoyment from the experience, and expressed stronger intentions to adopt the technology. This techno-optimism likely reflects both trust in centralized regulation and a collective orientation that foregrounds communal benefits over individual control. Such cultural alignment

with rule-based ethics and institutional authority has been found to correlate with higher baseline trust in AVs, especially when the technology is perceived as state-endorsed or norm-conforming (Awad et al., 2018). These patterns are consistent with findings from the Moral Machine Experiment, which observed that Eastern countries tend to emphasize social-order preservation and deontological logic rather than strictly utilitarian reasoning, especially in dilemmas involving law-abiding citizens and pedestrians.

By contrast, Mexican respondents—representing a transitional moral cluster—approached AVs with greater caution. They perceived higher risks and showed only moderate trust and enjoyment, translating into middling adoption intentions. This stance echoes the Moral Machine’s “Southern” cluster, where participants—primarily from Latin America—exhibited moderate utilitarian tendencies blended with additional social value judgments, such as prioritizing women and physically fit individuals (Awad et al., 2018). This hybrid moral logic may reflect the complex realities of emerging economies, where exposure to global technological narratives coincides with local constraints—such as infrastructural unreliability and institutional ambiguity—that shape perceptions of innovation as both opportunity and risk. These attitudes may mirror real-world concerns about uneven infrastructure, evolving safety standards, and economic constraints that temper enthusiasm for cutting-edge innovations. Yet their responses still exceeded U.S. levels on some well-being indices, suggesting an openness to new mobility solutions tempered by pragmatic considerations.

U.S. participants—anchored in a Western individualist tradition that emphasizes personal agency and skepticism toward centralized authority—reported the lowest trust and hedonic appeal, alongside moderate risk perceptions. Their comparatively muted adoption intentions reflect both a cultural premium on manual control and a more critical stance toward emergent automation. This mirrors Western countries’ responses in the Moral Machine data, where participants favored utilitarian logic in theory but also prioritized individual freedom and were less comfortable with abstract moral algorithms (Bonnefon et al., 2019; Awad et al., 2018). Although Western respondents often endorse saving the greatest number of lives in hypothetical scenarios, they frequently reject utilitarian configurations when imagining themselves or loved ones as passengers—highlighting a gap between moral ideals and real-world behavioral intentions (Bonnefon et al., 2016; Yoo et al., 2023).

Together, these descriptive findings validate our tri-national design: by drawing on empirically supported moral clusters identified in the Moral Machine’s global study of over 490,000 participants across 233 countries (Awad et al., 2018), we capture how moral worldviews, institutional trust, and lived infrastructures jointly influence the psychology of AV acceptance. This theoretical triangulation allows us to go beyond demographic generalizations and tap into deeper moral and emotional templates that shape AV-related judgments across cultures.

The structural models revealed a striking cultural divergence in how autonomy level shapes performance trust. In Mexico, higher autonomy translated directly into greater trust, whereas in both the U.S. and China this relationship was absent. This suggests that Mexican respondents—positioned between individualist skepticism and collectivist deference—may be particularly responsive to the technical sophistication of higher-level automation, forming a confidence bond not seen in either Western or East Asian contexts. This could reflect a heuristic association between technological complexity and perceived competence, particularly in societies where regulatory guidance may be inconsistent, and consumers rely on symbolic cues—like automation level—to evaluate safety.

When moral framing was introduced, only the Mexican sample showed a significant autonomy-by-morality interaction: here, utilitarian programming noticeably weakened the trust advantage of higher autonomy under mild accident scenarios. This aligns with existing research showing that people tend to find utilitarian programming unsettling in contexts involving potential personal sacrifice (Bonnefon et al., 2016;

Turpin et al., 2021), particularly when it threatens the safety of passengers. In transitional cultural contexts like Mexico, this unease may be amplified by the tension between public-oriented moral ideals and private risk aversion. In other words, while utilitarian logic may seem ethically sound in the abstract, it raises acute emotional discomfort when the possibility of personal harm becomes salient—even in hypothetical form. In contrast, neither the U.S. nor China showed a consistent trust response to ethical framing alone—suggesting that appeals to greater-good algorithms do not uniformly alter baseline confidence. These findings resonate with studies indicating that, despite general utilitarian endorsement, many individuals reject AV moral decisions that override emotional intuitions or imply arbitrary sacrifice (Zhu et al., 2022; Yoo et al., 2023).

Across all three countries, however, the full three-way interaction between autonomy level, moral logic, and accident severity consistently influenced performance trust—though the exact patterns varied by cultural context. In the Western samples, higher autonomy paired with self-protective logic increased trust when accidents were mild but reversed under high-severity conditions. Under utilitarian logic, the benefits of higher autonomy disappeared. In China, the pattern flipped: fully autonomous vehicles gained trust under mild utilitarian framing, but confidence dropped sharply in severe accident scenarios. These divergent shifts reflect how each culture balances personal safety, collective welfare, and emotional responses to catastrophic failures. They also illustrate the dynamic role of affective framing: when emotionally charged accident depictions are introduced, trust judgments shift dramatically—revealing that rational evaluations of automation can be quickly overruled by visceral reactions to imagined harm.

These findings are also in line with framing theory, which emphasizes how media narratives shape public interpretation and emotional reactions (Goffman, 1974; Kahneman and Tversky, 1979). In the case of AVs, emotionally charged coverage of fatal accidents can amplify perceived vulnerability and override trust, especially when the AV’s moral logic is seen as threatening (Shariff et al., 2017; Zhang et al., 2024). Negative or sensational media framing has been shown to distort public perception, erode confidence, and dampen adoption intentions even when statistical safety supports the technology (Cai et al., 2023). This reflects the availability heuristic, where vivid and emotionally salient events—like AV crashes—dominate memory and risk assessment, leading users to overestimate danger and underestimate long-term benefits. Such framing effects may be particularly potent in cultures where technological skepticism is already high or where users feel psychologically distant from the engineering and regulation of these systems.

Finally, we observed consistent downstream pathways from trust (and risk) to both hedonic well-being and behavioral intentions across all three countries. Greater performance trust enhanced users’ enjoyment and strengthened purchase intentions, while perceived risk diminished both outcomes. In Mexico and China, increased risk perceptions were particularly damaging to adoption intentions. These results reinforce prior research showing that trust is central to promoting hedonic well-being and behavioral uptake of AV technologies (Meyer-Waarden and Cloarec, 2022; Dogruel et al., 2015; Klobas et al., 2019). Perceived performance risk, in contrast, has been widely documented to erode user confidence, intensify technology anxiety, and reduce willingness to adopt (Featherman and Pavlou, 2003; Bertrandias et al., 2021; Hohenberger et al., 2016). This cascading effect—where reduced trust and heightened risk collectively diminish well-being and suppress behavioral intentions—highlights the centrality of emotional security in the AV adoption process. Consumers are not merely evaluating technological utility; they are negotiating complex feelings of control, responsibility, and vulnerability.

7.2. Practical contributions

Our Chinese sample entered the study with significantly higher

baseline trust, enjoyment, and adoption intentions—reflecting strong confidence in centralized regulation and a collective orientation toward technological stewardship. Managers and policymakers in similar collectivist markets should therefore foreground official endorsements and standardized safety credentials in all communications. For example, co-branding AV offerings with national or regional transport authorities, publishing aggregated performance metrics under government oversight, and engaging in visible collaborations with state-run infrastructure projects will reinforce the existing techno-optimism we observed. By tying product messaging directly to trusted institutions, stakeholders can amplify the communal benefits that Chinese participants intuitively associate with rule-conforming, state-backed innovations—thereby cementing high baseline trust before any ethical or accident framing is introduced.

Mexican respondents showed modest trust and enjoyment overall, but displayed a pronounced anxiety toward utilitarian programming in high-autonomy scenarios—where utilitarian logic significantly weakened the trust advantage of higher automation. This indicates that, in transitional moral climates, abstract appeals to “greater good” algorithms may inadvertently heighten emotional discomfort. To address this, AV providers should offer dual ethical-mode demonstrations, allowing users to experience both passenger-focused and communal-focused decision settings in mild-severity simulations. By presenting side-by-side case studies—such as short video vignettes comparing how each mode handles a low-impact obstacle—companies can help potential adopters concretely see that even utilitarian algorithms preserve passenger safety under realistic conditions. Grounding ethical choices in tangible, controlled scenarios directly responds to the autonomy-by-morality interaction we documented and helps bridge the gap between abstract moral endorsement and real-world comfort.

Across all three countries, performance trust and hedonic well-being were strongly enhanced by direct exposure, while perceived risk consistently undermined both enjoyment and intentions to adopt. This pattern underscores the power of the availability heuristic: vivid, emotionally salient crash narratives dominate risk assessments, but positive firsthand experiences can recalibrate those judgments. To capitalize on this, stakeholders should deploy scalable, low-barrier trial programs—such as short autonomous shuttle loops in urban centers, curated pilot-drive events in diverse traffic conditions, or subscription-based access to “shared AV pods.” Structured debriefings after each trial, highlighting real telemetry data on obstacle detection and emergency braking, will reinforce the safety narrative. By systematically replacing sensationalized accident imagery with personal, affect-rich memories of smooth, reliable rides, these experiential initiatives directly counteract the negative framing effects we found and lay the emotional groundwork for broader adoption.

7.3. Limitations and future research directions

Our study has several limitations that future research could address to enhance the robustness and generalizability of our findings. First, although our tri-national samples provide clear contrasts among Western individualist, Latin-American transitional, and East Asian collectivist contexts, they inevitably simplify the rich diversity within each region. By focusing solely on the United States, Mexico, and China, we capture prototypical moral-cognitive clusters but may miss nuanced subcultural or regional variations that influence AV attitudes. For example, urban versus rural residents in any of these countries may differ markedly in infrastructure exposure and risk tolerance—factors our balanced, quota-matched samples could not fully disentangle. Nevertheless, our results still robustly illustrate how broad cultural orientations shape baseline trust, moral-framing sensitivity, and autonomy-trust linkages.

Second, our reliance on hypothetical scenarios and media-style vignettes—while necessary for systematic manipulation—may limit ecological validity. Participants reacted to brief news articles and

written descriptions of SAE Levels 2 and 5, but these stimuli cannot fully replicate the visceral experience of riding in or observing an actual AV. Although our findings consistently show that direct exposure boosts performance trust and hedonic well-being, future field validations are needed to confirm that the autonomy-by-morality interactions and severity effects we observed manifest similarly in real-world test drives or pilot deployments.

Third, the binary framing of moral logic (self-protective vs. utilitarian) and autonomy level (low vs. high) omits the spectrum of intermediate ethical algorithms and SAE Levels 3–4 that now exist in many commercial vehicles. Our results revealed a unique Mexican sensitivity to utilitarian programming under high autonomy, but it remains unclear how more nuanced moral heuristics—or graduated autonomy tiers—would influence trust dynamics. While simplifying these dimensions sharpened our ability to detect clear cross-cultural patterns, future work should explore additional configurations to map the full landscape of AV moral design and user response.

Building on our demonstration of cultural divergence in autonomy-trust linkages, future studies should incorporate longitudinal designs that track how trust and risk perceptions evolve with repeated AV exposure. In particular, embedding surveys within ongoing pilot programs or ride-hailing partnerships could reveal whether the strong baseline techno-optimism in China or the transitional caution in Mexico attenuate or intensify over time, thereby testing the robustness of our cross-sectional findings.

To deepen understanding of moral framing effects beyond a binary choice, subsequent research might introduce hybrid or user-adjustable ethical settings—and measure not only stated preferences but also real-time decision behaviors under stress. For instance, by simulating variable moral algorithms in driving simulators or virtual reality environments, scholars can observe how affective reactions to “greater-good” logic shift when participants feel genuine immersion and agency, extending the autonomy-by-morality interactions we documented in our experiment.

Finally, expanding the cultural scope to include additional moral-cognitive clusters—such as South Asian collectivist or Northern European egalitarian societies—would test the generalizability of our tri-national design. Coupled with mixed-methods approaches (e.g., in-depth interviews, focus groups), this broader framework could uncover latent ethical considerations—like notions of fairness or dignity—that our current vignettes did not address. By triangulating across quantitative experiments, qualitative insights, and field trials, future research can build on our results to craft ethically attuned, culturally sensitive AV systems for the global marketplace.

CRedit authorship contribution statement

Thomas Teychenié: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Julien Cloarec:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lars Meyer-Waarden:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Adnan, N., Md Nordin, S., Bin Bahruddin, M.A., Ali, M., 2018. How trust can drive forward the user acceptance to the technology? In-vehicle technology for autonomous vehicle. *Transport. Res. Pol. Pract.* 118, 819–836. <https://doi.org/10.1016/j.tra.2018.10.019>.
- Ahluwalia, R., Burnkrant, R.E., Unnava, H.R., 2000. Consumer response to negative publicity: the moderating role of commitment. *J. Mark. Res.* 37 (2), 203–214. <https://doi.org/10.1509/jmkr.37.2.203.18734>.
- Asante, I.O., Jiang, Y., Miao, M., 2024. Exploring the motivating factors for using live-streaming and their influence on consumers' hedonic well-being: The mediating effect of psychological engagement. *Psychology & marketing* 41 (1), 27–44.
- Asgari, H., Jin, X., 2019. Incorporating attitudinal factors to examine adoption of and willingness to pay for autonomous vehicles. *Transp. Res. Rec.: J. Transport. Res. Board* 2673 (8), 418–429. <https://doi.org/10.1177/0361198119839987>.
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J., Rahwan, I., 2018. The moral machine experiment. *Nature*. <https://doi.org/10.1038/s41586-018-0637-6>.
- Bauer, J.J., McAdams, D.P., Sakaeda, A.R., 2005. Interpreting the good life: growth memories in the lives of mature, happy people. *J. Personality Soc. Psychol.* 88 (1), 203–217. <https://doi.org/10.1037/0022-3514.88.1.203>.
- BBC, 2023. Cruise driverless cars suspended after accident. <https://www.bbc.com/news/technology-67212093>.
- Bertrandias, L., Lowe, B., Sadik-Rozsnyai, O., Carricano, M., 2021. Delegating decision-making to autonomous products: a value model emphasizing the role of well-being. *Technol. Forecast. Soc. Change* 169, 120846. <https://doi.org/10.1016/j.techfore.2021.120846>.
- Bigman, Y.E., Gray, K., 2020. Life and death decisions of autonomous vehicles. *Nature* 579 (7797), E1–E2. <https://doi.org/10.1038/s41586-020-1987-4>.
- Bonnefon, J.-F., Rahwan, I., Shariff, A., 2024. The moral psychology of artificial intelligence. *Annu. Rev. Psychol.* 75 (1), 653–675. <https://doi.org/10.1146/annurev-psych-030123-113559>.
- Bonnefon, J.-F., Shariff, A., Rahwan, I., 2016. The social dilemma of autonomous vehicles. *Science* 352 (6293), 1573–1576. <https://doi.org/10.1126/science.aaf2654>.
- Bonnefon, J.-F., Shariff, A., Rahwan, I., 2019. The trolley, the bull bar, and why engineers should care about the ethics of autonomous cars [point of view]. *Proc. IEEE* 107 (3), 502–504. <https://doi.org/10.1109/JPROC.2019.2897447>.
- Borah, A., Tellis, G.J., 2016. Halo (Spillover) effects in social media: do product recalls of one brand hurt or help rival brands? *J. Mark. Res.* 53 (2), 143–160. <https://doi.org/10.1509/jmr.13.0009>.
- Bruno, G., Spoto, A., Sarlo, M., Lotto, L., Marson, A., Cellini, N., Cutini, S., 2024. Moral reasoning behind the veil of ignorance: an investigation into perspective-taking accessibility in the context of autonomous vehicles. *Br. J. Psychol.* 115 (1), 90–114. <https://doi.org/10.1111/bjop.12679>.
- Cai, Y., Jing, P., Wang, B., Jiang, C., Wang, Y., 2023. How does “over-hype” lead to public misconceptions about autonomous vehicles? A new insight applying causal inference. *Transport. Res. Pol. Pract.* 175, 103757. <https://doi.org/10.1016/j.tra.2023.103757>.
- Carter, O.B.J., Loft, S., Visser, T.A.W., 2024. Meaningful communication but not superficial anthropomorphism facilitates human-automation trust calibration: the human-automation trust expectation model (HATEM). *Hum. Factors: The Journal of the Human Factors and Ergonomics Society* 66 (11), 2485–2502. <https://doi.org/10.1177/00187208231218156>.
- Chen, Y., Ghosh, M., Liu, Y., Zhao, L., 2019. Media coverage of climate change and sustainable product consumption: evidence from the hybrid vehicle market. *J. Mark. Res.* 56 (6), 995–1011. <https://doi.org/10.1177/0022243719865898>.
- Chen, Z., Jin, J., Li, M., 2022. Does media coverage influence firm green innovation? The moderating role of regional environment. *Technol. Soc.* 70, 102006. <https://doi.org/10.1016/j.techsoc.2022.102006>.
- CNBC, 2025. Waymo, Toyota strike partnership to bring self-driving tech to personal vehicles. CNBC. <https://www.cnbc.com/2025/04/29/waymo-toyota-partner-to-bring-self-driving-tech-to-personal-vehicles.html>.
- Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 13 (3), 319. <https://doi.org/10.2307/249008>.
- Dietvorst, B.J., Simmons, J.P., Massey, C., 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* 144 (1), 114–126. <https://doi.org/10.1037/xge0000033>.
- Dogruel, L., Joeckel, S., Bowman, N.D., 2015. The use and acceptance of new media entertainment technology by elderly users: development of an expanded technology acceptance model. *Behav. Inf. Technol.* 34 (11), 1052–1063. <https://doi.org/10.1080/0144929X.2015.1077890>.
- Dong, M., Bonnefon, J.-F., Rahwan, I., 2024. Toward human-centered AI management: methodological challenges and future directions. *Technovation* 131, 102953. <https://doi.org/10.1016/j.technovation.2024.102953>.
- Dos Santos, F.L.M., Duboz, A., Grosso, M., Raposo, M.A., Krause, J., Mourtzouchou, A., Balahur, A., Ciuffo, B., 2022. An acceptance divergence? Media, citizens and policy perspectives on autonomous cars in the European Union. *Transport. Res. Pol. Pract.* 158, 224–238. <https://doi.org/10.1016/j.tra.2022.02.013>.
- Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P., 2003. The role of trust in automation reliance. *Int. J. Hum. Comput. Stud.* 58 (6), 697–718. [https://doi.org/10.1016/S1071-5819\(03\)00038-7](https://doi.org/10.1016/S1071-5819(03)00038-7).
- Everett, J.A.C., Colombatto, C., Awad, E., Boggio, P., Bos, B., Brady, W.J., Chawla, M., Chituc, V., Chung, D., Drupp, M.A., Goel, S., Grosskopf, B., Hjorth, F., Ji, A., Kealoha, C., Kim, J.S., Lin, Y., Ma, Y., Maréchal, M.A., et al., 2021. Moral dilemmas and trust in leaders during a global health crisis. *Nat. Hum. Behav.* 5 (8), 1074–1088. <https://doi.org/10.1038/s41562-021-01156-y>.
- Everett, J.A.C., Pizarro, D.A., Crockett, M.J., 2016. Inference of trustworthiness from intuitive moral judgments. *J. Exp. Psychol. Gen.* 145 (6), 772–787. <https://doi.org/10.1037/xge0000165>.
- Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transport. Res. Pol. Pract.* 77, 167–181. <https://doi.org/10.1016/j.tra.2015.04.003>.
- Featherman, M., Jia, S. (Jasper), Calif, C.B., Hajli, N., 2021. The impact of new technologies on consumers beliefs: reducing the perceived risks of electric vehicle adoption. *Technol. Forecast. Soc. Change* 169, 120847. <https://doi.org/10.1016/j.techfore.2021.120847>.
- Featherman, M.S., Pavlou, P.A., 2003. Predicting e-services adoption: a perceived risk facets perspective. *Int. J. Hum. Comput. Stud.* 59 (4), 451–474. [https://doi.org/10.1016/S1071-5819\(03\)00111-3](https://doi.org/10.1016/S1071-5819(03)00111-3).
- Financial Times, 2025. China's Baidu in Talks to Launch Robotaxis in Europe. *Financial Times*. <https://www.ft.com/content/27775dcd-1a61-4562-b60f-f3d331e25032>.
- Gill, T., 2020. Blame it on the self-driving car: how autonomous vehicles can alter consumer morality. *J. Consum. Res.* 47 (2), 272–291. <https://doi.org/10.1093/jcr/ucaa018>.
- Goffman, E., 1974. *Frame Analysis: an Essay on the Organization of Experience*. Harvard University Press.
- Graf-Vlachy, L., König, A., Banfield, R., Rauch, M., Boutalikkakis, A., 2023. The innovator's media dilemma: how journalists cover incumbents' adoption of discontinuous technologies. *J. Prod. Innovat. Manag.* 40 (1), 3–29. <https://doi.org/10.1111/jpim.12651>.
- Han, L., Sun, R., Gao, F., Zhou, Y., Jou, M., 2019. The effect of negative energy news on social trust and helping behavior. *Comput. Hum. Behav.* 92, 128–138. <https://doi.org/10.1016/j.chb.2018.11.012>.
- Harman, D., 1967. A single factor test of common method variance. *Journal of Psychology* 35 (1967), 359–378.
- Hayes, A.F., 2022. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, Third edition. The Guilford Press.
- He, X., Happee, R., Wang, M., 2024. A new computational perceived risk model for automated vehicles based on potential collision avoidance difficulty (PCAD). *Transport. Res. C Emerg. Technol.* 166, 104751. <https://doi.org/10.1016/j.trc.2024.104751>.
- Higgins, E.T., 1997. Beyond pleasure and pain. *Am. Psychol.* 52 (12), 1280–1300. <https://doi.org/10.1037/0003-066X.52.12.1280>.
- Hohenberger, C., Spörrle, M., Welpe, I.M., 2016. How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups. *Transport. Res. Pol. Pract.* 94, 374–385. <https://doi.org/10.1016/j.tra.2016.09.022>.
- Jaspers, E.D.T., Pearson, E., 2022. Consumers' acceptance of domestic Internet-of-Things: the role of trust and privacy concerns. *J. Bus. Res.* 142, 255–265. <https://doi.org/10.1016/j.jbusres.2021.12.043>.
- Jobst, L.J., Bader, M., Moshagen, M., 2023. A tutorial on assessing statistical power and determining sample size for structural equation models. *Psychological Methods* 28 (1), 207.
- Joo, Y.K., Kim, B., 2023. Selfish but socially approved: the effects of perceived collision algorithms and social approval on attitudes toward autonomous vehicles. *Int. J. Hum. Comput. Interact.* 39 (19), 3717–3727. <https://doi.org/10.1080/10447318.2022.2102716>.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.
- Kenešei, Z., Ásványi, K., Kökény, L., Jászberényi, M., Miskolczi, M., Gyulavári, T., Syahrivar, J., 2022. Trust and perceived risk: how different manifestations affect the adoption of autonomous vehicles. *Transport. Res. Pol. Pract.* 164, 379–393. <https://doi.org/10.1016/j.tra.2022.08.022>.
- Klobas, J.E., McGill, T., Wang, X., 2019. How perceived security risk affects intention to use smart home devices: a reasoned action explanation. *Comput. Secur.* 87, 101571. <https://doi.org/10.1016/j.cose.2019.101571>.
- Koopman, P., Wagner, M., 2017. Autonomous vehicle safety: an interdisciplinary challenge. *IEEE Intelligent Transportation Systems Magazine* 9 (1), 90–96. <https://doi.org/10.1109/ITS.2016.2583491>.
- Korneeva, E., Salge, T.O., Teubner, T., Antons, D., 2023. Tracing the legitimacy of Artificial Intelligence: a longitudinal analysis of media discourse. *Technol. Forecast. Soc. Change* 192, 122467. <https://doi.org/10.1016/j.techfore.2023.122467>.
- Lankton, N., McKnight, D.H., State University, Michigan, Tripp, J., Baylor University, 2015. Technology, humanness, and trust: rethinking trust in technology. *J. Assoc. Inf. Syst. Online* 16 (10), 880–918. <https://doi.org/10.17705/1jais.00411>.
- Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. *Hum. Factors: The Journal of the Human Factors and Ergonomics Society* 46 (1), 50–80. <https://doi.org/10.1518/hfes.46.1.50.30392>.
- Li, Y., Cassard, M., Holmes, B., Wu, H., 2025. Variations in media framing of movements in China, France, and the U.S.: an intersectional approach. *Br. J. Sociol.* 76 (1), 114–135. <https://doi.org/10.1111/1468-4446.13153>.
- Liang, C., Song, Y., Wang, B., 2022. The influence of social function of Mobile game augmented reality and virtual reality environment on Mobile phone users' addiction. *Wireless Commun. Mobile Comput.* 2022. <https://doi.org/10.1155/2022/2903124>.
- Lijarcio, I., Useche, S.A., Llamazares, J., Montoro, L., 2019. Perceived benefits and constraints in vehicle automation: data to assess the relationship between driver's features and their attitudes towards autonomous vehicles. *Data Brief* 27, 104662. <https://doi.org/10.1016/j.dib.2019.104662>.
- Mayer, M.M., Bell, R., Buchner, A., 2021. Self-protective and self-sacrificing preferences of pedestrians and passengers in moral dilemmas involving autonomous vehicles. *PLoS One* 16 (12), e0261673. <https://doi.org/10.1371/journal.pone.0261673>.

- McKinsey, 2024. Autonomous Vehicles Moving Forward: Perspectives from Industry Leaders. McKinsey & Company. <https://www.mckinsey.com/features/mckinsey-center-for-future-mobility/our-insights/autonomous-vehicles-moving-forward-perspectives-from-industry-leaders>.
- McKnight, D.H., Carter, M., Thatcher, J.B., Clay, P.F., 2011. Trust in a specific technology: an investigation of its components and measures. *ACM Transactions on Management Information Systems* 2 (2), 1–25. <https://doi.org/10.1145/1985347.1985353>.
- McKnight, D.H., Chervany, N.L., 2001. What trust means in E-Commerce customer relationships: an interdisciplinary conceptual typology. *Int. J. Electron. Commer.* 6 (2), 35–59. <https://doi.org/10.1080/10864415.2001.11044235>.
- Meyer-Waarden, L., Cloarec, J., 2022. “Baby, you can drive my car”: psychological antecedents that drive consumers’ adoption of AI-powered autonomous vehicles. *Technovation* 109, 102348. <https://doi.org/10.1016/j.technovation.2021.102348>.
- Mirzahosseini, H., Mashhadloo, M., 2024. Modeling the effect of autonomous vehicles (AVs) on the accessibility of the transportation network. *Scientific reports* 14 (1), 9292.
- Moshagen, M., Erdfelder, E., 2016. A new strategy for testing structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal* 23 (1), 54–60.
- Myers, S., Everett, J.A.C., 2025. People expect artificial moral advisors to be more utilitarian and distrust utilitarian moral advisors. *Cognition* 256, 106028. <https://doi.org/10.1016/j.cognition.2024.106028>.
- Nair, G.S., Bhat, C.R., 2021. Sharing the road with autonomous vehicles: perceived safety and regulatory preferences. *Transport. Res. C Emerg. Technol.* 122, 102885. <https://doi.org/10.1016/j.trc.2020.102885>.
- NBC, 2025. Waymo passenger records dizzying ride as car drives in circles. <https://www.nbcnews.com/news/us-news/watch-waymo-driverless-robotaxi-sends-airport-passenger-spinning-circle-rcna186672>.
- Newsweek, 2025. Chinese driverless car chief breaks silence after first fatal crash. <https://www.newsweek.com/chinese-driverless-car-chief-breaks-silence-after-first-fatal-crash-2054333>.
- Nordhoff, S., Stapel, J., He, X., Gentner, A., Happee, R., 2021. Perceived safety and trust in SAE level 2 partially automated cars: results from an online questionnaire. *PLoS One* 16 (12), e0260953. <https://doi.org/10.1371/journal.pone.0260953>.
- Noy, I.Y., Shinar, D., Horrey, W.J., 2018. Automated driving: safety blind spots. *Saf. Sci.* 102, 68–78. <https://doi.org/10.1016/j.ssci.2017.07.018>.
- Oliver, A.G., Campbell, R., Graffin, S., Bundy, J., 2023. Media coverage of earnings announcements: how newsworthiness shapes media volume and tone. *J. Manag.* 49 (4), 1213–1245. <https://doi.org/10.1177/01492063221080125>.
- Osburg, V.-S., Yoganathan, V., Kunz, W.H., Tarba, S., 2022. Can (A)I give you a ride? Development and validation of the CRUISE framework for autonomous vehicle services. *J. Serv. Res.* 25 (4), 630–648. <https://doi.org/10.1177/10946705221118233>.
- Ostrom, A.L., Fotheringham, D., Bitner, M.J., 2019. Customer acceptance of AI in service encounters: understanding antecedents and consequences. In: Maglio, P.P., Kieliszewski, C.A., Spohrer, J.C., Lyons, K., Patrício, L., Sawatani, Y. (Eds.), *Handbook of Service Science*, II. Springer International Publishing, pp. 77–103. https://doi.org/10.1007/978-3-319-98512-1_5.
- Othman, K., 2021. Public acceptance and perception of autonomous vehicles: a comprehensive review. *AI and Ethics*. <https://doi.org/10.1007/s43681-021-00041-8>.
- Pagiavlas, S., Kalaigianam, K., Gill, M., Bliese, P.D., 2022. Regulating product recall compliance in the digital age: evidence from the “Safe Cars Save Lives” campaign. *J. Market.* 86 (5), 135–152. <https://doi.org/10.1177/00222429211023016>.
- Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2000. A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybern. Syst. Hum.* 30 (3), 286–297. <https://doi.org/10.1109/3468.844354>.
- Pavlou, P.A., 2003. Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. *Int. J. Electron. Commer.* 7 (3), 101–134. <https://doi.org/10.1080/10864415.2003.11044275>.
- Penmetasa, P., Adanu, E.K., Wood, D., Wang, T., Jones, S.L., 2019. Perceptions and expectations of autonomous vehicles – a snapshot of vulnerable road user opinion. *Technol. Forecast. Soc. Change* 143, 9–13. <https://doi.org/10.1016/j.techfore.2019.02.010>.
- Perello-March, J., Burns, C.G., Woodman, R., Birrell, S., Elliott, M.T., 2024. How do drivers perceive risks during automated driving scenarios? An fNIRS neuroimaging study. *Hum. Factors: The Journal of the Human Factors and Ergonomics Society* 66 (9), 2244–2263. <https://doi.org/10.1177/00187208231185705>.
- Pizzi, G., Scarpì, D., 2020. Privacy threats with retail technologies: a consumer perspective. *J. Retailing Consum. Serv.* 56, 102160. <https://doi.org/10.1016/j.jretconser.2020.102160>.
- Pizzi, G., Scarpì, D., Pichierrì, M., Vannucci, V., 2019. Virtual reality, real reactions?: comparing consumers’ perceptions and shopping orientation across physical and virtual-reality retail stores. *Comput. Hum. Behav.* 96, 1–12. <https://doi.org/10.1016/j.chb.2019.02.008>.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology* 88 (5), 879.
- Ram, S., Sheth, J.N., 1989. Consumer resistance to innovations: the marketing problem and its solutions. *J. Consum. Market.* 6 (2), 5–14. <https://doi.org/10.1108/EUM000000002542>.
- Reuters, 2024. CEO Barra backs GM’s push for autonomous vehicles. <https://www.reuters.com/business/autos-transportation/ceo-barra-backs-gms-push-autonomous-vehicles-2024-05-23/>.
- Reuters, 2025. Italy tests self-driving vehicle in car sharing pilot. <https://www.reuters.com/business/autos-transportation/italy-tests-self-driving-vehicle-car-sharing-pilot-2025-01-22/>.
- Rittenberg, B.S.P., Holland, C.W., Barnhart, G.E., Gaudreau, S.M., Neyedli, H.F., 2024. Trust with increasing and decreasing reliability. *Hum. Factors: The Journal of the Human Factors and Ergonomics Society* 66 (12), 2569–2589. <https://doi.org/10.1177/00187208241228636>.
- Rosen, L.D., Weil, M.M., 1994. What we have learned from a decade of research (1983–1993) on “the psychological impact of technology.”. *ACM SIGCAS Computers and Society* 24 (1), 3–9.
- SAE, 2021. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles – SAE.
- Sharif, A., Bonnefon, J.-F., Rahwan, I., 2017. Psychological roadblocks to the adoption of self-driving vehicles. *Nat. Hum. Behav.* 1 (10), 694–696. <https://doi.org/10.1038/s41562-017-0202-6>.
- Singh, K., Grewal, R., 2023. Lobbying and product recalls: a study of the U.S. automobile industry. *J. Mark. Res.* 60 (4), 728–749. <https://doi.org/10.1177/00222437221131568>.
- Sinkovics, R.R., Stöttinger, B., Schlegelmilch, B.B., Ram, S., 2002. Reluctance to use technology-related products: development of a technophobia scale: reluctance to use technology-related products. *Thunderbird Int. Bus. Rev.* 44 (4), 477–494. <https://doi.org/10.1002/tie.10033>.
- Slovic, P., 1987. Perception of risk. *Science* 236 (4799), 280–285. <https://doi.org/10.1126/science.3563507>.
- Statista, 2025. Autonomous vehicles worldwide. <https://www.statista.com/topics/3573/autonomous-vehicle-technology/#topicOverview>.
- Steg, L., 2005. Car use: lust and must. Instrumental, symbolic and affective motives for car use. *Transport. Res. Pol. Pract.* 39 (2–3), 147–162. <https://doi.org/10.1016/j.tra.2004.07.001>.
- Subero-Navarro, Á., Pelegrín-Borondo, J., Reinares-Lara, E., Olarte-Pascual, C., 2022. Proposal for modeling social robot acceptance by retail customers: CAN model + technophobia. *J. Retailing Consum. Serv.* 64, 102813. <https://doi.org/10.1016/j.jretconser.2021.102813>.
- Swartz, D.G., Strand, I.E., 1981. Avoidance costs associated with imperfect information: the case of kepone. *Land Econ.* 57 (2), 139. <https://doi.org/10.2307/3145782>.
- Takaguchi, K., Kappes, A., Yearsley, J.M., Sawai, T., Wilkinson, D.J.C., Savulescu, J., 2022. Personal ethical settings for driverless cars and the utility paradox: an ethical analysis of public attitudes in UK and Japan. *PLoS One* 17 (11), e0275812. <https://doi.org/10.1371/journal.pone.0275812>.
- Thrash, T.M., Elliot, A.J., Maruskin, L.A., Cassidy, S.E., 2010. Inspiration and the promotion of well-being: tests of causality and mediation. *J. Personality Soc. Psychol.* 98 (3), 488–506. <https://doi.org/10.1037/a0017906>.
- Turpin, M.H., Walker, A.C., Fugelsang, J.A., Sorokowski, P., Grossmann, I., Bialek, M., 2021. The search for predictable moral partners: predictability and moral (character) preferences. *J. Exp. Soc. Psychol.* 97, 104196. <https://doi.org/10.1016/j.jesp.2021.104196>.
- Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. *Cogn. Psychol.* 5 (2), 207–232. [https://doi.org/10.1016/0010-0285\(73\)90033-9](https://doi.org/10.1016/0010-0285(73)90033-9).
- Umar, Z., Jareño, F., González, M.D.L.O., 2021. The impact of COVID-19-related media coverage on the return and volatility connectedness of cryptocurrencies and fiat currencies. *Technol. Forecast. Soc. Change* 172, 121025. <https://doi.org/10.1016/j.techfore.2021.121025>.
- Van Brummelen, J., O’Brien, M., Gruyer, D., Najjaran, H., 2018. Autonomous vehicle perception: the technology of today and tomorrow. *Transport. Res. C Emerg. Technol.* 89, 384–406. <https://doi.org/10.1016/j.trc.2018.02.012>.
- Venkatesh, Thong, Xu, 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q.* 36 (1), 157. <https://doi.org/10.2307/41410412>.
- Victor, T.W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., Ljung Aust, M., 2018. Automation expectation mismatch: incorrect prediction despite eyes on threat and hands on wheel. *Hum. Factors: The Journal of the Human Factors and Ergonomics Society* 60 (8), 1095–1116. <https://doi.org/10.1177/0018720818788164>.
- Wang, S., Jiang, Z., Noland, R.B., Mondschein, A.S., 2020. Attitudes towards privately-owned and shared autonomous vehicles. *Transport. Res. F Traffic Psychol. Behav.* 72, 297–306. <https://doi.org/10.1016/j.trf.2020.05.014>.
- Waterman, A.S., 2008. Reconsidering happiness: a eudaimonist’s perspective. *J. Posit. Psychol.* 3 (4), 234–252. <https://doi.org/10.1080/17439760802303002>.
- Wilde, G.J., Claxton-Oldfield, S.P., Platenius, P.H., 1985. Risk homeostasis in an experimental context. In: *Human Behavior and Traffic Safety*. Springer US, Boston, MA, pp. 119–149.
- Wilson, C., Gyi, D., Morris, A., Bateman, R., Tanaka, H., 2022. Non-driving related tasks and journey types for future autonomous vehicle owners. *Transport. Res. F Traffic Psychol. Behav.* 85, 150–160. <https://doi.org/10.1016/j.trf.2022.01.004>.
- Wirtz, J., Patterson, P.G., Kunz, W.H., Gruber, T., Lu, V.N., Paluch, S., Martins, A., 2018. Brave new world: service robots in the frontline. *J. Serv. Manag.* 29 (5), 907–931. <https://doi.org/10.1108/JOSM-04-2018-0119>.
- Wotton, M.E.L., Bennett, J.M., Modesto, O., Challinor, K.L., Prabhakaran, P., 2022. Attention all ‘drivers’: you could be to blame, no matter your behaviour or the level of vehicle automation. *Transport. Res. F Traffic Psychol. Behav.* 87, 219–235. <https://doi.org/10.1016/j.trf.2022.03.021>.
- Wu, M., Wang, N., Yuen, K.F., 2023. Can autonomy level and anthropomorphic characteristics affect public acceptance and trust towards shared autonomous vehicles? *Technol. Forecast. Soc. Change* 189, 122384. <https://doi.org/10.1016/j.techfore.2023.122384>.

- Yokoi, R., Nakayachi, K., 2021. Trust in autonomous cars: exploring the role of shared moral values, reasoning, and emotion in safety-critical decisions. *Human factors* 63 (8), 1465–1484.
- Yoo, S., Kumagai, J., Morita, T., Park, Y.G., Managi, S., 2023. Who to sacrifice? Modeling the driver's dilemma. *Transport. Res. Pol. Pract.* 178, 103872. <https://doi.org/10.1016/j.tra.2023.103872>.
- Zhang, Q., Wallbridge, C.D., Jones, D.M., Morgan, P.L., 2024. Public perception of autonomous vehicle capability determines judgment of blame and trust in road traffic accidents. *Transport. Res. Pol. Pract.* 179, 103887. <https://doi.org/10.1016/j.tra.2023.103887>.
- Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R., Zhang, W., 2019. The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transport. Res. C Emerg. Technol.* 98, 207–220. <https://doi.org/10.1016/j.trc.2018.11.018>.
- Zhu, A., Yang, S., Chen, Y., Xing, C., 2022. A moral decision-making study of autonomous vehicles: expertise predicts a preference for algorithms in dilemmas. *Pers. Individ. Differ.* 186, 111356. <https://doi.org/10.1016/j.paid.2021.111356>.