Technology in Society 87 (2026) 103389



Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc





The human factor in AI governance: Dispositional predictors of public concern and regulatory support☆

Dayeon Eom [a], Julianne Renner [a], Yijia Erika Zhu [a], Anqi Shao [a,c], Soobin Choi [b,d], Todd Newman [a], Dominique Brossard [a,b], Dietram Scheufele [a,b,*]

[a] Department of Life Sciences Communication, University of Wisconsin-Madison, United States
[b] Morgridge Institute for Research, United States
[c] Global Institute of Sustainability and Innovation, Arizona State University, United States
[d] Division of Communication & Media, Ewha Women's University, South Korea

ARTICLE INFO

Keywords:
Artificial intelligence
Authenticity
Agency
Context
Regulation
Algorithm aversion

ABSTRACT

This study uses a 4 × 2 factorial experiment ($N = 1{,}640$) to examine how perceptions of AI as autonomous agents versus decision-support tools, along with application context (medical diagnosis, screenwriting, law enforcement, or genetic screening), influence public concerns about AI and support for its regulation among U.S. adults. Counter to theoretical expectations based on algorithm aversion literature, neither AI agency nor application context produced significant main effects on public attitudes. In contrast, three key dispositions consistently predicted outcomes across contexts: authenticity preference, concern about disproportionate impacts on vulnerable groups, and general attitudes toward AI. These effects remained consistent across all experimental conditions, indicating that public responses to AI are fundamentally driven by pre-existing individual dispositions rather than contextual features or implementation approaches. These results are consistent with moral foundations and mind-sponge accounts of value-based judgment. The findings imply that risk-tiered regulatory frameworks, while necessary, are unlikely to be sufficient on their own; durable governance will require deliberative engagement with the authenticity and equity intuitions that drive public response, and longitudinal attention to how those orientations are themselves reshaped by sustained algorithmic exposure.

Artificial Intelligence (AI) applications span everyday activities, from composing emails and personalized news recommendations to powering facial recognition on video doorbells, as well as more complex systems such as autonomous vehicles, medical diagnostic tools, and law enforcement technologies (Choi et al., 2025; Reis et al., 2024). With the permeation of AI in daily life, there are advantages such as convenience and efficiency, but it also raises questions about social and ethical implications. For example, AI technologies in law enforcement, such as facial recognition software that performs poorly at recognizing faces of people of color (Allyn, 2020) and predictive policing algorithms that have relied on biased historical arrest data (Verma, 2022), have raised important equity concerns. These issues have prompted scrutiny (Farooq, 2024; MacMillan et al., 2025), and continuous efforts are being made to call for improved fairness and accountability in the development and deployment of such systems. As AI is rapidly changing and reshaping how humans and modern societies function, there is a growing research need for a more granular understanding of public perceptions of the technology.

Recent public opinion polls indicate that public concern about AI is rising in tandem with increased exposure and adoption (Bao et al., 2022; Dreksler et al., 2025). A majority of Americans report feeling more concerned than excited about AI in daily life, citing potential threats to privacy, employment, and human relationships (Tyson & Kikuchi, 2023). In addition to the practical concerns, discomfort with AI may stem from its perceived agency, particularly when it mimics human decision-making in emotionally or ethically charged contexts. Human-like capabilities, imbued with a perception of self-directed goals or intentions, can evoke the "uncanny valley" effect, where AI triggers

☆ All authors have agreed to this submission, and the article is not currently being considered for publication by any other print or electronic journal. Correspondence concerning this article should be addressed to Dietram Scheufele, Life Sciences Communication Department, University of Wisconsin-Madison, 1545 Observatory Drive, Madison, WI 53706.

* Corresponding author. Department of Life Sciences Communication, University of Wisconsin-Madison, United States.
E-mail addresses: deom4@wisc.edu (D. Eom), scheufele@wisc.edu (D. Scheufele).

https://doi.org/10.1016/j.techsoc.2026.103389
Received 2 October 2025; Received in revised form 5 May 2026; Accepted 7 May 2026
Available online 11 May 2026
0160-791X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

feelings of unease or distrust (Seyama & Nagayama, 2007; Sundar, 2020). Such negative sensation intensifies when AI is perceived as making or influencing high-stakes decisions. For instance, in 2023, 60 percent of U.S. adults reported being uncomfortable with medical doctors relying on AI for diagnosis or treatment recommendations (Tyson et al., 2023). Similar concerns surfaced in the contexts of determining eligibility for social services and for rendering verdicts in criminal trials (Loewen et al., 2024). These responses reflect not just anxiety about AI's performance, but deeper unease about delegating moral and interpretive agency to non-human systems (Kieslich et al., 2022).

Despite growing scholarly attention to public attitudes toward AI, existing research has predominantly focused on contextual factors such as application demand (subjective vs. objective task orientation) and task characteristics (moral vs. non-moral decision-making) as primary drivers (Castelo et al., 2019; Van Berkel et al., 2022). While these studies have documented how different contexts influence algorithm aversion or appreciation, few have systematically examined how perceived AI agency, the degree to which systems operate autonomously versus under human supervision, interacts with contextual factors to shape public attitudes. Furthermore, limited research has systematically examined whether these contextual and agential effects are actually meaningful compared to stable individual differences in dispositions. Recent work has begun identifying individual predictors such as neuroticism and familiarity with AI, but few studies have directly compared the importance of dispositional versus situational factors in shaping AI attitudes (Mahmud et al., 2022). This represents a gap in understanding the psychological foundations of public AI governance preferences, as current theoretical frameworks may underestimate the role of pre-existing individual dispositional predictors, such as preference for authenticity and concern about social justice, factors that could have implications for AI regulation and public engagement.

To address this research gap, we conducted an online survey experiment that directly tested the relative influence of contextual factors versus individual dispositions in shaping public responses to AI implementation. Rather than assuming contextual effects dominate, our study systematically compared how application contexts and AI agencies influence public attitudes against the predictive power of stable individual differences in dispositions. Our study contributes to the literature in two key ways: 1) it tests whether application contexts and AI agency, independently or combined, meaningfully differentiate public concerns about AI or support for its regulation, and 2) it directly compares the explanatory power of experimental manipulations against individual dispositions. By quantifying the situational versus dispositional factors, the study provides empirically grounded recommendations for effective public engagement and policy development approaches that acknowledge the primacy of individual dispositions in governance preferences.

1. Artificial Intelligence as a wicked problem

Artificial intelligence refers to “computer systems that perform tasks or make decisions that usually require human intelligence” (Zhang & Dafoe, 2019). Public awareness of AI has risen markedly in recent years, spurred by the widespread adoption of applications such as OpenAI's ChatGPT (Tyson & Kikuchi, 2023). As a result, most individuals now interact with AI systems either knowingly or passively in their daily lives (Faverio & Kikuchi, 2026; Maese, 2025). This rapid and ubiquitous integration of AI has prompted growing public interest and concern (Tyson & Kikuchi, 2023). While attitudes toward AI may shift as the technology evolves, examining current public hesitations is essential for guiding ethical development, governance, and communication strategies. Public opposition to technology serves as an early warning signal of potential social frictions, which should be addressed through democratic deliberation. This ensures that regulatory responses and design choices align with public values, rather than lagging behind them.

AI governance can be understood through the lens of wicked problems, which are defined as complex policy challenges that require societies to negotiate difficult trade-offs among competing priorities and value systems without straightforward solutions (Strauβ, 2021). Like climate change or poverty, regulating AI entails characteristics of “wicked science”: issues that are technically intricate, socially contested, and in need of timely action before risks escalate beyond existing frameworks (Austin & Háji, 2023). Unlike other technologies that individuals or communities may choose to disengage from, AI increasingly permeates critical infrastructures, making opting out practically infeasible. Compounding these structural barriers are three interlocking moral dilemmas that, drawing from evolutionary psychology, help explain why governance so often stalls not from a lack of awareness but from a failure to convert awareness into coordinated action (Ho & Ho, 2025). First, the *tragedy of the commons* captures how short-term individual gains, such as rushing undertested AI products to market, erode collective welfare over time. Second, the *tragedy of commonsense morality* describes how divergent normative frameworks across stakeholder groups make shared governance difficult to construct. Third, the *tragedy of apathy* reflects how people tend to favor inaction over change and to discount future risks in favor of present comfort, producing decision-paralysis rather than urgent action, even among those who recognize AI's risks (Austin & Háji, 2023).

The current state of public trust and expert opinion regarding AI oversight compounds these governance challenges. Recent polls suggest that while the public and AI experts diverge in their enthusiasm and predictions for AI, they converge on a desire for stronger oversight and personal control (McClain et al., 2025). Roughly six in ten U.S. adults and 56% of surveyed experts report greater concern about under-regulation than over-regulation of AI, reflecting limited trust in the government's capacity to manage this technology effectively (McClain et al., 2025). This lack of trust highlights a critical gap in the accountability mechanisms necessary to safeguard against AI misuse. Moreover, public skepticism toward AI governance is reinforced by broader patterns of institutional distrust across sectors where AI will be deployed, including healthcare, media, and the legal system (Robinson, 2020). Tailored regulation could therefore serve a dual purpose: mitigating AI-related risks while rebuilding the institutional credibility necessary for effective oversight (Fournier-Tombs, 2021).

Due to its complexity, addressing AI as a wicked problem requires a multi-stakeholder perspective that treats the public not as an obstacle to innovation, but as a partner in identifying where collective benefits lie (Elia & Margherita, 2018; Hess, 2015). Such collaboration necessitates structured processes to weigh competing values and craft workable solutions. Crucially, meaningful engagement depends first on understanding the public's concerns and the evaluative frameworks through which they interpret AI (Hess, 2015). Furthermore, for AI governance to be both effective and ethically grounded, it must be guided by frameworks that are conceptually robust enough to capture the normative complexity of emerging technologies while remaining operationally feasible within public sector institutions (Wilson & Van Der Velden, 2022). Many previous attempts at AI governance have faced limitations because they either lacked conceptual clarity or proved impractical to implement at scale, particularly in bureaucratic or resource-constrained contexts. Bridging this gap between ethical aspiration and policy implementation remains a central challenge in responsible AI governance.

2. Contextual and agential factors in AI attitudes

Previous research has found that users’ attitudes and policy support are shaped in specific use scenarios, such as daily life tasks or enjoyment. Specifically, task-oriented scenarios have been found to generate more positive attitudes, while job replacement scenarios have created negative responses (Yan et al., 2024). Additionally, evaluations of the costs and benefits of AI, as well as the resulting tensions between feeling served (personal and societal benefits) and feeling exploited (perceived

costs to the individual and society), also result in differences across contexts (Dorotic et al., 2024). For example, facial recognition technology may be accepted for use in schools to fortify security measures, but not accepted in public spaces due to privacy concerns (Choung et al., 2024). Similarly, a systematic review of 85 studies across autonomous driving, healthcare, and robotics reveals that acceptance factors vary by application context (Ismatullaev & Kim, 2024). While privacy risks emerge as critical barriers to the adoption of autonomous vehicles and healthcare AI due to safety and medical data concerns, these same concerns have a minimal impact in robotics, where users already perceive these systems as beneficial (Ismatullaev & Kim, 2024).

The importance of context-specificity is echoed in policy debates. Effective AI regulation must be tailored not only to the technology's capabilities but also to the values of its intended audience and the potential severity of its consequences (Beaudouin et al., 2020). Context-sensitive governance is particularly important as different AI applications vary not only in their technical operation but also in their symbolic implications, as some challenge long-standing professional norms (e.g., patient-provider relationship) or touch on emotionally charged decisions. Additionally, regulatory approaches must consider whether AI systems are deployed for localized or global purposes, as this shapes the distribution of benefits and harms (Beaudouin et al., 2020).

The European Union's AI Act serves as a key regulatory example that governs AI based on both risk assessment and specific use cases (European Parliament, 2023). High-risk contexts under the Act include AI systems embedded in products covered by the EU's product safety legislation, such as toys, aviation, automobiles, and medical devices, as well as AI applications in sensitive areas like law enforcement and migration control, which must be registered in an EU-wide database (European Parliament, 2023). Drawing on these guidelines, we selected healthcare diagnosis, law enforcement, and genetic screening as high-risk contexts. For comparison, we selected screenwriting as a low-risk context. Creative applications like content generation typically fall outside the Act's high-risk categories and face fewer regulatory requirements due to their limited potential for direct harm to individuals' safety, rights, or fundamental freedoms. This contrast allows us to examine how different risk classifications affect perceived concerns about AI and support for its regulation.

Notably, this study examines public attitudes toward AI at the level at which citizens typically encounter governance questions, as a general-purpose technology deployed across varying application contexts, rather than as technically distinct system architectures (e.g., generative models, recommender algorithms, embodied robots). This framing is deliberate: public discourse and regulatory frameworks alike tend to organize AI not by underlying technical architecture but by application context and associated risk (Dreksler et al., 2025; European Parliament, 2023). Table 1 summarizes the four experimental contexts along dimensions most relevant to governance, including risk classification, public familiarity, example applications, and relevant existing policies.

H1a. High-risk contexts (healthcare, law enforcement, genetic screening) will elicit greater perceived concerns than low-risk contexts (screenwriting).

H1b. High-risk contexts will elicit greater support for regulation than low-risk contexts.

Beyond context, AI's perceived agency, how much control or autonomy it appears to have in human-computer interaction, is a key factor shaping public reactions, particularly when systems are deployed in consequential contexts (Jiang et al., 2024). AI agencies exist on a spectrum, from fully autonomous systems such as self-driving vehicles to decision support tools that operate under human oversight. As perceived AI agency increases, especially in high-stakes applications, public acceptance tends to decline (Jiang et al., 2024; Wu et al., 2021). Conversely, when AI is portrayed as augmenting rather than replacing human decision-making, it is viewed more positively and as more trustworthy (Li et al., 2024). Previous research has shown that the type of human-AI collaboration and users' outcome expectations interact to influence consumer evaluations and intentions to use AI products and services (Yue & Li, 2023).

This pattern of declining acceptance with increased AI agency aligns with broader findings in algorithm aversion research, where humans often reject algorithmic decision-making despite algorithms' superior performance across many tasks. Early studies in this area focused primarily on contexts where algorithms had demonstrably failed (Dietvorst et al., 2015), but subsequent research found that algorithm aversion can emerge even before users witness any algorithmic errors, indicating that rejection is not solely driven by observed poor performance (Longoni et al., 2019). Building on these findings, researchers have developed more comprehensive frameworks identifying specific factors that contribute to algorithm aversion, including algorithm characteristics such as agency, performance, capabilities, and degree of human involvement, as well as human agent characteristics like expertise and social distance (Jussupow et al., 2020). Among these factors, agency deserves particular attention because it represents a distinctive feature that sets AI apart from traditional technologies and fundamentally shapes how users perceive and interact with these systems.

At its core, assigning high agency to AI in ethically sensitive contexts challenges social norms about which decisions should remain human, rooted in empathy, lived experience, and social responsibility (Federspiel et al., 2023). These perceived threats to human judgment and moral boundaries intensify concerns about AI, especially when it is seen as encroaching on roles that are central to professional identity or cultural priorities. Thus, agency functions not only as a technical feature but also as a psychological and symbolic cue, shaping whether people resist or support AI, particularly through their judgments about the threats it poses to uniquely human decision making.

Table 1
Overview of AI application contexts examined in this study.

Context	Risk Level (per EU AI Act)	Domain	Key Governance Concern	Example AI Applications	Relevant Policies in the U.S.[a]
Medical Diagnosis	High	Healthcare	Patient safety, liability	Diagnostic imaging AI, clinical decision support	California's AB 489; Colorado's SB 205; FDA regulation of AI/ML-enabled medical devices
Law Enforcement	High	Criminal justice	Bias, civil liberties	Facial recognition, predictive policing	Montana Code Annotated § 44-15-106; Maryland SB 182/ HB 338
Genetic Screening	High	Bioethics	Privacy, discrimination	Polygenic risk scoring, prenatal screening	Genetic Information Nondiscrimination Act (GINA); Genetic Information Privacy Act (GIPA)
Screenwriting	Low	Creative	Authorship, labor	Content generation	U.S. Copyright Office Report on AI and Copyright; WGA Minimum Basic Agreement

[a] *Note*. The U.S. currently lacks comprehensive federal AI legislation. The policies listed here represent a fragmented landscape of domain-specific regulations, state-level laws, and industry agreements rather than a unified governance framework. Moreover, several of these policies predate the widespread integration of AI into their respective domains. For instance, GINA (2008) prohibits discrimination based on genetic information in health insurance and employment but does not explicitly address AI-driven applications such as polygenic risk scoring or algorithmic prenatal screening. Similarly, law enforcement AI oversight relies primarily on a patchwork of state and municipal laws rather than federal standards.

H2a. Greater AI agency will increase perceived concerns.

H2b. Greater AI agency will increase support for regulation.

RQ1. How does the interaction between application context and AI agency influence 1) perceived concerns about AI implementation and 2) support for AI regulation?

3. Dispositions against AI

While context and agency provide valuable insights into algorithm aversion, the complexity of this phenomenon extends beyond these structural factors alone. Recent studies suggest that algorithm aversion stems from multiple sources, including individual-specific factors such as psychological characteristics, demographic variables, and personal familiarity with both algorithms and specific tasks (Mahmud et al., 2022). The widespread aversion to algorithms in morally charged decision-making contexts appears particularly multifaceted (Jauernig et al., 2022), indicating that resistance may be driven as much by dispositions as by the technical characteristics of the systems themselves. To better understand these individual differences in algorithm aversion, we incorporate the following dispositional factors into our analysis.

3.1. General attitude toward AI

An individual's attitude toward AI encompasses cognitive evaluations (e.g., reliability, safety, usefulness), affective reactions (e.g., fear, excitement, anxiety), and ethical judgments (e.g., transparency, fairness, accountability), forming stable evaluative frameworks that consistently shape responses to AI across contexts rather than fleeting opinions (Acosta-Enriquez et al., 2024; Lahusen et al., 2024; Schepman & Rodway, 2020). These attitudinal dispositions function as interpretive filters, where negative attitudes toward emerging technology reduce support for its use across applications and amplify perceptions of threat, particularly for high-agency systems in morally sensitive contexts. Such negative effects are demonstrated through strong public support for human supervision requirements and temporal stability in attitude-behavior relationships (Bao et al., 2022; Hemesath & Tepe, 2024; Kim et al., 2014; Korukonda, 2005). As the public becomes increasingly familiar with AI technologies over time through growing media attention (Ouchchy et al., 2020), individuals are likely to develop more crystallized attitudes toward AI (Bao et al., 2022). These attitudes serve as baseline considerations that shape how people later evaluate specific AI implementations and regulatory preferences. In this way, pre-existing attitudes become essential to consider when examining patterns of concern and support.

3.2. Authenticity Preference

With AI systems becoming increasingly human-like, the concept of authenticity has become central to understanding what people value and perceive as trustworthy in human-AI interactions, with key dimensions including transparency, alignment with users' dispositions and identities, and perceived sincerity (Vo et al., 2025). This study conceptualizes authenticity preference as the degree to which individuals prioritize transparency and human-likeness when interacting with AI agents, building on prior psychology research showing that people generally prefer authentic experiences across domains from food consumption to interpersonal relationships (Newman, 2019). This tendency may also extend to human-machine interactions, where strong authenticity preferences may reduce receptivity to AI agents. Studies have documented how perceived lack of transparency can drive rejection of algorithmic systems, often stemming from concerns about the "black box" nature of AI decision-making, where users cannot understand or trace the reasoning behind algorithmic outputs (Castelo et al., 2019). The black box problem becomes particularly pronounced when AI systems make consequential decisions without providing clear explanations, creating a fundamental mismatch between users' expectations for understandable processes and the opaque nature of many algorithmic systems. This suggests that authenticity preferences may represent a key psychological mechanism underlying algorithm aversion, where rejection stems not merely from performance concerns but from fundamental incompatibility between users' dispositions about decision-making and the perceived inscrutability of algorithmic systems.

3.3. Concerns about disproportionate impact

Automated decisions can systematically disadvantage marginalized groups across contexts, from credit scoring to medical evaluation and hiring, with inequities often rooted in biased training data and learned patterns that contribute to digital discrimination (Benjamin, 2019; Ferrer et al., 2021; Roselli et al., 2019). Many scholars argue that bias is not a technical glitch to be fixed but a reflection of broader societal inequities embedded within AI systems, raising concerns that some level of disproportionate impact may be structurally inescapable (Benjamin, 2019). While public awareness of algorithmic bias has grown in recent years, technical solutions continue to lag behind these fairness concerns (Faverio & Kikuchi, 2026; Belenguer, 2022; Nazer et al., 2023). Research demonstrates that concerns about disproportionate impact significantly influence algorithm acceptance in morally contested scenarios. For instance, studies examining life-and-death decisions such as ventilator allocation during COVID-19 found low public preference for algorithmic decision-making (Filiz et al., 2023). As AI systems become increasingly prevalent in decision-making, understanding how concerns about disproportionate impact shape public acceptance becomes essential for both technology development and governance strategies.

Together, these constructs represent distinct but complementary dimensions of public evaluation, technological orientation, moral-psychological comfort, and social justice concerns that offer a more holistic understanding of how people form attitudes toward AI implementation and regulation. This study addresses the following research hypotheses.

H3a. Negative attitudes toward AI will increase concerns and support for regulation.

H3b. Stronger authenticity preferences will increase concerns and support for regulation.

H3c. Greater perception of disproportionate impact will increase concerns and support for regulation.

4. Methods

4.1. Sample and data description

In March 2025, we recruited U.S. adults using online panels from Forthright (Bovitz, Inc.), with survey programming via Qualtrics. Participation was voluntary, with informed consent required. Quality control excluded participants who completed the survey in under 4 min (30% of the median completion time) (Greszki et al., 2015), sped through consecutive grid questions, had completion rates below 70%, or failed attention checks. Minimal missing values (less than 1%) were addressed using Multivariate Imputation by Chained Equations (MICE) in R with Predictive Mean Matching (Van Buuren & Groothuis-Oudshoorn, 2011). Of the 2037 individuals who began the survey, 1640 met the criteria (80.51%). A power analysis was conducted using G*Power 3.1 to determine the required sample size. For detecting a small effect size ($f = 0.10$) with 95% power using analysis of variance at $\alpha = 0.05$, the analysis indicated a minimum required sample size of $N = 1501$. Our sample size of $N = 1640$ exceeded this requirement, ensuring adequate statistical power.

The final sample consisted of 1640 participants, including 813 females (49.6%), 794 males (48.4%), and 33 non-binary individuals

(2.0%). The average age was 47.85 years (*SD* = 16.66; *median* = 47, *range* = 19–88). White/Caucasian respondents comprised the largest ethnic group (68%), followed by Black or African American (14%) and Asian American (3%). In terms of education, 25% reported some college credit without a degree, 24% had completed high school, 23% held a bachelor's degree, and 10% reported having postgraduate degrees.

4.2. Experiment design

This study employed a 4 (medical diagnosis vs. screenwriting vs. law enforcement vs. genetic screening) × 2 (AI-driven vs. human-led) plus control factorial design (See Fig. 1 for a conceptual diagram of the experiment design). After consenting to participate, participants answered questions about the moral decision-making ability of AI; they were then randomly assigned to one of nine experimental conditions, each featuring a mockup online news article with varying combinations of images and texts (See Fig. 2, as well as the *Stimuli* section below). Post treatment, they answered questions about perceived concerns about AI, support for regulation, authenticity preference, perception of disproportionate impact, general attitude toward AI, and basic demographic information.

Multiple AI image generation systems, including DALL·E, MidJourney, and ChatGPT, were used to create high-quality visual stimuli. Outputs from MidJourney were ultimately selected for their realistic image generation capabilities that responded effectively to minor prompt modifications. To inform stimulus design, relevant AI news articles were analyzed to guide the visual presentation. The stimuli were designed using four contexts (medical diagnosis, law enforcement, screenwriting, and genetic screening) under two conditions: AI-agency focus and human-agency focus. These contexts were selected to represent diverse applications with varying levels of risk, societal impact, public familiarity, and expertise requirements (technical vs. creative), allowing us to test whether public perceptions of AI are context-dependent.

The same AI-generated image was used across both agency conditions, but was edited differently to emphasize which entity was central to the message. In each agency condition, the highlighted entity (either AI or human) was made more prominent as a decision authority by adding a shadow over the non-highlighted entity. For example, in the medical diagnosis scenario, when using a human-agency focus, doctors were highlighted in the image to demonstrate that AI was being used as an assistant to humans. Meanwhile, for the AI-agency focus, the AI system the doctors were interacting with was highlighted instead. Along with visual cues, textual variations in stimuli highlighted the entity with agency in each context. For instance, headlines for human-agency focused conditions read, "Doctors Harness AI to Enhance Medical Diagnoses." Meanwhile, the AI-agency focused condition read, "AI Automates Medical Diagnoses." The subheadings and image captions were similarly manipulated to ensure that the agency was clear.

The mockup online articles were tailored to mimic content from the *Calgary Herald*, a Canadian news publication expected to be unfamiliar to the US-based participants recruited for the survey. To enhance authenticity and control for confounding variables, all posts displayed the publication date to replicate typical online news article formatting. All experimental materials were reviewed by a panel of science communication researchers to confirm they accurately represented the intended conditions.

To assess whether participants perceived the experimental scenarios as intended, we included manipulation-check items asking them to identify both the application context and whether the decision-maker was an AI system or a human. Overall, 81% of participants answered both items correctly. Failures likely reflect a mix of satisficing and confusion, as the agency check used a seven-point scale ranging from "Only humans" to "Only AI." Although this design captured nuanced perceptions of human-AI involvement, it may also have introduced ambiguity, particularly for participants choosing midpoint options (e.g., equal human and AI involvement). However, the overall pass rate is consistent with levels reported in prior survey-experimental research (Oppenheimer et al., 2009). We retained all participants in the main analyses to preserve the full range of responses and maintain ecological validity, recognizing that in real-world settings, individuals may not always attend closely to system details. Supplementary analyses excluding manipulation-check failures yielded substantively similar results (see Tables 2 and 3 in Supplementary Appendix), indicating that the findings are not driven by misperceptions of the experimental manipulations.

4.3. Measures

Perceived Concerns about AI. We adapted thirteen items from Esmaeilzadeh et al. (2021) to measure perceived concerns about AI on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*) (Esmaeilzadeh et al., 2021). Items included: "AI systems may produce inaccurate results," "Reliability of AI is questionable," "AI poses new legal liabilities," "Tracing mistakes in AI is challenging," "Data might be shared without consent," "It's unclear who is accountable for AI errors," and "AI's lack of emotional intelligence may diminish empathy," among others. The scores from these thirteen items were averaged ($M = 5.481$, $SD = 0.922$; Cronbach's $\alpha = 0.914$), with higher scores indicating greater concerns about AI.

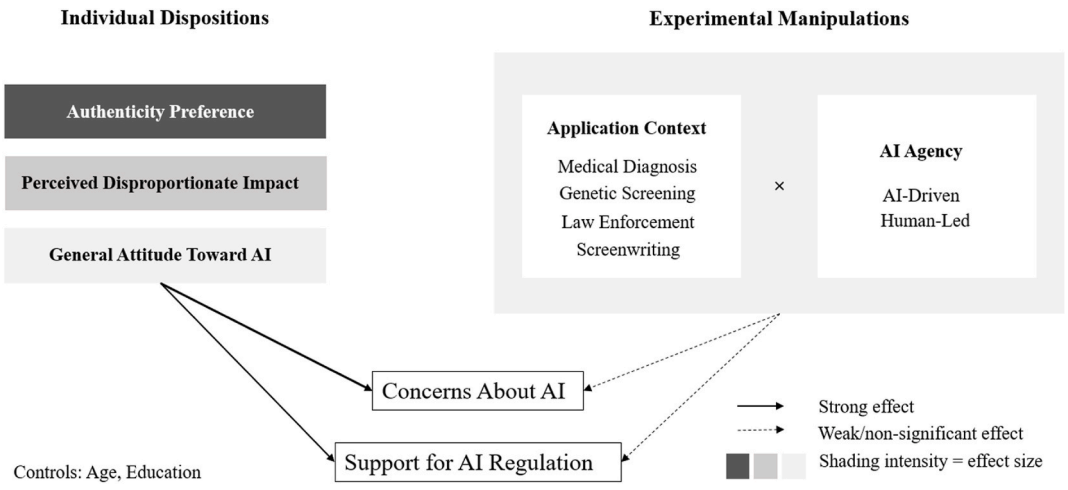


Fig. 1. Conceptual Diagram of Experiment Design
Note. A control condition with no contextual framing was also included in the experimental design, but is omitted from this diagram for visual clarity.



Fig. 2. Stimuli Under Different Experimental Conditions
Note. From Left to Right: AI-Agency Focused vs. Human-Agency Focused.
Top to Bottom: Medical Diagnosis, Genetic Screening, Screenwriting, Law Enforcement.

Table 2
Analysis of variance (ANOVA) results predicting concerns about AI.

Source	Sum of Squares	*df*	Mean of Squares	*F*	*P*	η_p^2
Agency	0.570	1	0.570	1.410	0.235	<0.001
Context	0.213	3	0.071	0.175	0.913	<0.001
Agency × Context	3.461	3	1.154	2.854	0.036*	0.006
Authenticity Preference	539.941	1	539.941	1335.684	<0.001***	0.480
Perceived Disproportionate Impact	78.797	1	78.797	194.924	<0.001***	0.119
General Attitude Toward AI	15.685	1	15.685	38.800	<0.001***	0.026
Age	0.075	1	0.075	0.186	0.666	0.000
Education	6.661	1	6.661	16.479	<0.001***	0.011
Residuals	584.535	1446	0.404			

Note. $*p < 0.05$ $**p < 0.01$ $***p < 0.001$.

Table 3
Analysis of variance (ANOVA) results predicting support for AI regulation.

Source	Sum of Squares	*df*	Mean of Squares	*F*	*P*	η_p^2
Agency	0.437	1	0.437	0.393	0.531	<0.001
Context	4.447	3	1.482	1.332	0.262	0.003
Agency × Context	9.541	3	3.180	2.857	0.036*	0.006
Authenticity Preference	475.299	1	475.299	427.043	<0.001***	0.228
Perceived Disproportionate Impact	34.626	1	34.626	31.111	<0.001***	0.021
General Attitude Toward AI	4.464	1	4.464	4.011	0.045*	0.003
Age	0.596	1	0.596	0.536	0.464	<0.001
Education	0.021	1	0.021	0.019	0.890	<0.001
Residuals	1609.400	1446	1.113			

Note. $*p < 0.05$ $**p < 0.01$ $***p < 0.001$.

Support **for Regulation.** We used a single item to assess the intention to regulate AI (Faverio & Kikuchi, 2026). Participants rated their agreement with the statement "As AI applications become more widespread, to what extent do you agree or disagree with the following statement? Artificial intelligence should be regulated" on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*). Higher scores indicated stronger support for AI regulation ($M = 6.015$, $SD = 1.236$).

Authenticity Preference. Inspired by the multi-dimensional approaches to authenticity measurement in marketing and consumer research literature (Bui et al., 2024; Lee & Kim, 2024; Yang & Battocchio, 2021), we developed six items to assess authenticity preference in AI-generated content. Items were measured on a 7-point scale (1 = *strongly disagree* to 7 = *strongly agree*) and included: "The use of AI should always be clearly labeled as such," "I believe that AI can be as authentic as humans" (reverse coded), "The use of AI diminishes the authenticity of the final product," "I feel that entities should be transparent about the extent of AI involvement," "AI lacks the personal touch that makes humans more authentic," and "I am more likely to trust humans compared to AI." Higher scores indicated a higher authenticity preference ($M = 5.507$, $SD = 0.948$; Cronbach's $\alpha = 0.764$) (See Table 6 in Supplementary Appendix for factor loadings).

Perceived Disproportionate Impact. Perceptions of the disproportionate impact of AI were assessed using a single item adapted from the Yale Program on Climate Change Communication's Climate Change in the American Mind: Climate Justice survey (Carman et al., 2023). The original item asked whether respondents think global warming harms some groups of people in the U.S. more than others. We adapted this item for the AI context: "Do you think the negative impacts of generative AI will harm some groups of people in the United States more than others?" Participants responded on a 5-point scale (1 = *definitely not* to 5 = *definitely*). Higher scores indicated a stronger belief in the disproportionate impacts of AI ($M = 3.707$, $SD = 1.009$).

General Attitude Toward AI. We used eight semantic differential-type items to assess general attitudes toward AI (Kim & Jung, 2023). Participants rated AI on a 7-point scale between the following bipolar adjectives: Negative to Positive, Untrustworthy to Trustworthy, Harmful to Beneficial, Scary to Exciting, Hard to understand to Easy to understand, Unethical to Ethical, Unsafe to Safe, and Biased to Objective. The scores from these eight items were averaged ($M = 4.055$, $SD = 1.294$; Cronbach's $\alpha = 0.925$), with higher scores indicating more positive attitudes toward AI.

Demographic Controls. Demographic Controls. We tested several demographic variables as potential covariates, including age, education, and gender. Age and education were retained as control variables because they showed significant relationships with the outcome variables and because prior research has established that both factors are consistently associated with differences in technology perceptions and risk sensitivity (Ismatullaev & Kim, 2024; Czaja et al., 2006). Although some prior research has identified gender differences in AI attitudes (Beets et al., 2023; Sartori & Bocca, 2023), one-way ANOVAs revealed no significant differences across experimental conditions in the present study, and gender was therefore not retained as a covariate. Including these demographic controls allowed us to isolate the effects of the primary predictors while accounting for potential confounding influences from baseline differences in digital literacy, information processing, or structural access to AI-related experiences.

4.4. Analytical approaches

The hypotheses and research questions were tested using analysis of covariance (ANCOVA). Upon checking assumptions, we conducted one-way and factorial ANOVA tests to assess the impact of independent variables and moderators on the dependent variable. Additionally, to understand complex three-way interactions in the overall model, we conducted subgroup analyses by context and ran separate ANOVA models for each application context. Results are reported with complete test statistics, including sums of squares, degrees of freedom, mean squares, F-values, effect sizes, and corresponding p-values, based on a predetermined significance threshold of $\alpha = 0.05$. Lastly, to ensure the validity of our moderators and confirm they were not themselves influenced by the experimental manipulation, we conducted preliminary ANOVA tests examining the effects of message conditions on each moderator variable. Results indicated that none of the moderators, authenticity preference, perceived disproportionate impact, and general attitude toward AI, were significantly affected by the message conditions (all $ps > 0.05$). This suggests that the moderators are statistically independent of the manipulated message conditions.

5. Results

5.1. Main effects

Analysis revealed no significant main effects for the experimental manipulations on concerns about AI, indicating that neither how AI systems are framed (as AI-driven versus human-led) nor the specific application context influences public worry (see Table 2). The agency manipulation showed no significant effect, $F(1,1446) = 1.41$, $p = 0.235$, $\eta_p^2 < 0.001$, meaning that participants expressed statistically negligible differences in concern levels. This finding contradicts H2a, which predicted that higher AI agency would increase concern levels. The hypothesis that autonomous AI systems would generate greater worry compared to human-supervised systems received no empirical support. Similarly, application context failed to reach significance, $F(3,1446) = 0.18$, $p = 0.913$, $\eta_p^2 < 0.001$, demonstrating that concern levels remained remarkably consistent whether AI was described in medical diagnosis, law enforcement, genetic screening, or screenwriting contexts. This result fails to support H1a, which predicted that high-risk contexts (medical, legal, genetic screening) would elicit greater concern than lower-risk contexts (screenwriting). Contrary to expectations, participants showed equivalent concern levels across all application domains.

Regulatory support showed the same pattern of minimal experimental influence (see Table 3). Agency framing produced no significant effect, $F(1,1446) = 0.39$, $p = 0.531$, $\eta_p^2 < 0.001$, indicating that whether AI systems are presented as autonomous versus human-supervised tools does not meaningfully shift public support for government oversight. This result provides no support for H2b's prediction that higher AI agency would increase support for regulation. Application context similarly failed to predict regulatory preferences, $F(3,1446) = 1.33$, $p = 0.262$, $\eta_p^2 = 0.003$. H1b, which predicted that high-risk contexts would elicit greater support for regulation, was not supported.

5.2. Dispositional predictors

Authenticity preference emerged as the dominant predictor of both outcomes, demonstrating the powerful role of individual disposition in shaping AI attitudes. For concerns about AI implementation, $F(1,1446) = 1335.68$, $p < 0.001$, $\eta_p^2 = 0.480$, the substantial F-statistic reflects the strength and consistency of this relationship across participants. For regulatory support, authenticity preferences showed $F(1,1446) = 427.04$, $p < 0.001$, $\eta_p^2 = 0.228$. These findings strongly support the authenticity preference prediction of H3b. Individuals with stronger preferences for human authenticity showed significantly higher AI concerns and greater regulatory support.

Perceived disproportionate impact served as another major predictor. For AI concerns, $F(1,1446) = 194.92$, $p < 0.001$, $\eta_p^2 = 0.119$. This disposition explains 12% of the variance, a medium to large effect, indicating that individuals who worry about AI exacerbating inequalities consistently express higher overall AI concerns. For regulatory support, perceived disproportionate impact concerns showed $F(1,1446) = 31.11$, $p < 0.001$, $\eta_p^2 = 0.021$, a smaller but still significant effect explaining 2% of the variance. This suggests that while concerns about inequitable AI impacts strongly predict general AI anxiety, they have a more modest (though still meaningful) relationship with regulatory preferences. People who worry about AI harming marginalized groups consistently favor oversight, but this connection is less pronounced than for overall concern levels. These results provide strong support for H3c.

The general attitude toward AI showed expected relationships with both outcomes. More negative general attitudes predicted higher concern levels, $F(1, 1446) = 38.80$, $p < 0.001$, $\eta_p^2 = 0.026$, explaining 3% of the variance. While this represents a small effect, it demonstrates that baseline technology attitudes influence domain-specific concerns in predictable ways. For regulatory support, the relationship was weaker but still significant, $F(1, 1446) = 4.01$, $p = 0.045$, $\eta_p^2 = 0.003$, explaining less than 1% of the variance. This minimal effect suggests that general AI attitudes have limited influence on regulatory preferences once other dispositions are considered, indicating that governance support is driven more by specific dispositions (authenticity, fairness) than by general technology sentiment. These findings provide support for H3a, though the regulatory effect was smaller than concerns about AI.

5.3. Interaction effects

Both dependent variables showed a significant but negligible agency × context interaction. For concerns about AI, $F(3,1446) = 2.85$, $p = 0.036$, $\eta_p^2 = 0.006$, and for regulatory support, $F(3,1446) = 2.86$, $p = 0.036$, $\eta_p^2 = 0.006$. While achieving statistical significance due to our large sample size, these interactions explain only 0.6% of the variance in each outcome, below conventional thresholds for meaningful effects. These findings provide minimal support for RQ1, which asked how the interaction between application context and AI agency influences perceived concerns and regulatory support.

Post-hoc analyses revealed that for concerns about AI, genetic screening showed significantly greater concern for AI framing compared to human framing (estimate = 0.139, $p = 0.037$), while screenwriting demonstrated a marginally significant trend in the opposite direction, with human framing generating slightly more concern (estimate = -0.124, $p = 0.062$); medical and legal contexts showed no agency differences ($ps > 0.47$). For regulatory support, genetic screening again showed significance for AI framing (estimate = 0.240, $p = 0.029$), while law showed a marginally significant preference for human framing (estimate = -0.213, $p = 0.056$), with medical and screenwriting contexts showing no meaningful differences ($ps > 0.69$). These opposing patterns suggest context-specific responses to AI implementation approaches, though the minimal overall effect sizes indicate these contextual differences have limited practical influence (See Fig. 3).

Additional analyses examined three-way interactions between agency, context, and each dispositional predictor to assess whether individual differences moderate the agency × context effects (See Tables 4 and 5 in Supplementary Appendix). For AI concerns, only the agency × context × perceived disproportionate impact interaction reached significance, $F(3, 1425) = 3.13$, $p = 0.025$, while interactions with authenticity preference and general AI attitudes were non-significant ($ps > 0.84$). For regulatory support, both the agency × context × authenticity preference interaction, $F(3, 1425) = 4.13$, $p = 0.006$, and the agency × context × general AI attitude interaction, $F(3, 1425) = 3.71$, $p = 0.011$, achieved significance, while the perceived disproportionate impact interaction was non-significant ($p = 0.183$). However, these three-way interactions explained minimal additional variance beyond the substantial main effects of the dispositional predictors, suggesting that while individual differences may subtly alter how people respond to specific agency-context combinations, these moderating effects are of limited practical significance.

6. Discussion

6.1. Scientific and theoretical implications

The experimental hypotheses regarding the main effects of application context and AI agency were not supported. Contrary to predictions derived from algorithm aversion research, neither high-stakes contexts such as medical diagnosis and law enforcement nor differences between AI agency versus human-agency systems yielded meaningful differences in public responses. These null results are noteworthy given the emphasis in prior scholarship on task stakes and perceived machine autonomy as key drivers of opposition. Instead, these null effects raise the possibility that public resistance to AI is not easily shifted by contextual or technical cues, at least within the contexts examined here.

Nonetheless, the significant interaction between agency and context revealed patterns worth noting. Participants expressed greater concern and stronger support for regulation when autonomous AI was applied to

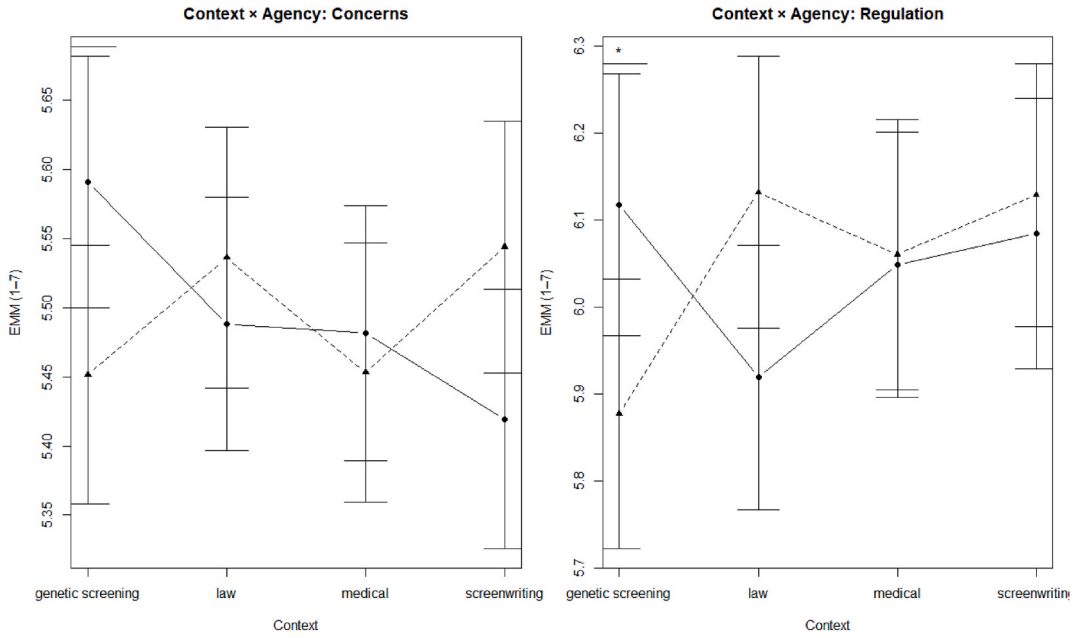


Fig. 3. Context and Agency Interaction Effects on Concerns and Regulation Support
Notes. Circles (solid lines) represent AI agency; triangles (dashed lines) represent human agency. Error bars show 95% confidence intervals. Contexts include genetic screening, law enforcement, medical diagnosis, and screenwriting.

genetic screening, while the reverse occurred in law enforcement, where human involvement elicited more concern. Although modest in magnitude, these effects suggest that agency perceptions are filtered through the cultural meanings of particular contexts. In genetic screening, where decisions implicate personal dignity, family life, and moral boundaries, autonomous AI may be perceived as overstepping the role of its judgment. In contrast, law enforcement raised sharper questions of accountability and legitimacy, with participants displaying more concern when human actors were responsible, likely reflecting public awareness of systemic bias and displaying machine heuristic (Sundar & Kim, 2019).

Even so, these context-dependent dynamics were relatively small compared to the role of individual dispositions. Authenticity preference alone explained nearly half of the variance in concerns about AI and almost a quarter of the variance in regulatory support, while perceived disproportionate impacts on vulnerable groups contributed additional explanatory power. That authenticity preference so powerfully predicts both concern and regulatory support suggests that the public may already be registering normative threats to qualities they regard as irreducibly human, and that these evaluations remain stable regardless of how AI is framed. By systematically comparing the explanatory weight of experimental manipulations with dispositional predictors, this study addresses a gap in the literature: individual dispositions explain substantially more variance in concerns and regulatory preferences than do situational cues.

These findings carry broader theoretical implications for AI governance. The persistence of dispositional effects indicates a form of attitudinal stability: attitudes toward emerging technologies are anchored in durable value orientations rather than easily swayed by contextual cues. Moral Foundations Theory holds that moral judgment is rooted in evolutionarily shaped intuitions that function as fast, automatic responses to perceived violations (Haidt & Joseph, 2004; Graham et al., 2013). Our results are consistent with this framework: rather than appearing to assess concern based on whether AI is deployed in healthcare versus law enforcement, or operates autonomously versus under human supervision, individuals appear to apply stable moral templates anchored in authenticity preferences and equity concerns that function as evaluative heuristics across contexts. This inertia helps explain why experimental manipulations produced only modest shifts and speaks directly to arguments that governance must begin with societal values and public expectations rather than treating them as downstream concerns (Janssen, 2025). For governance frameworks to be legitimate and resilient, they must engage the public not merely as end-users of AI but as stakeholders whose ethical orientations shape both acceptance and resistance.

Yet these value commitments should not be mistaken for permanent fixtures. Granular interaction thinking theory conceptualizes values as informational quanta continuously refined through interactions with surrounding environments, including algorithmic ones (Vuong et al., 2025). From this perspective, the dispositional stability observed here may represent a cross-sectional snapshot of orientations that are themselves in slow but consequential motion. Scholars of national and global value systems have further warned that sustained exposure to algorithmic media is accelerating a broader societal drift toward individualism and value fragmentation, gradually eroding the collective orientations on which legitimate governance dialogue depends (Stückelberger, 2024; Sweet, 2023). This creates a unique challenge: the stakeholder values that governance frameworks seek to engage are being reshaped by the very technologies under governance. Tracking how algorithmic interactions reconfigure these orientations over time, rather than only measuring them as they currently stand, represents one of the most pressing directions for future research.

The variability and contestation around these normative concerns reinforce AI's status as a wicked problem. As with other wicked issues such as climate change or gene editing, governance is unlikely to achieve consensus through technical fixes alone, and ignoring normative pluralities risks eroding trust and legitimacy. These findings also resonate with the mind-sponge framework (Vuong & Napier, 2015; Ho et al., 2022), which emphasizes how individuals filter new information through cost-benefit judgments informed by their dispositions. Applying this lens to AI governance highlights why human-centered regulation must look beyond technical performance metrics to the underlying motivations that shape acceptance and resistance. The strong predictive role of perceived disproportionate impact further converges with scholarship arguing that AI's moral risks concentrate among already marginalized populations (Benjamin, 2019; Valčo, 2024). Our finding

that this perception predicts both heightened concern and stronger regulatory preferences independently of application context adds to this literature by demonstrating that distributive justice orientations function as stable evaluative lenses not only in contexts with documented disparities but uniformly across AI applications.

6.2. Policy and practical implications

These theoretical insights carry direct implications for regulatory design, particularly in interpreting the modest yet significant interaction effects between context and agency. While these results offer partial support for tiered, risk-based regulatory models such as the EU AI Act, they also suggest limits to technocratic sorting. Differentiating regulation by application-specific risk is consistent with these results, but technocratic sorting alone may be insufficient to address diverse publics. Public receptivity depends not only on which contexts are regulated but also on how agency and risk are understood in tandem, particularly how responsibility and power are distributed within each context and who is benefiting. Responsible governance frameworks increasingly call for shifting from internal organizational priorities to broader societal considerations (Janssen, 2025; Taeihagh, 2021), and our findings offer empirical grounding for why this reorientation is necessary. When individual dispositions, however gradually evolving, consistently outperform contextual and technical cues in shaping public response, governance frameworks anchored primarily in technical risk classification may fail to capture the normative concerns that drive public acceptance and resistance. Rather than anchoring governance in rapidly shifting technical capabilities, centering the value orientations and normative expectations that publics bring to these debates, while building in mechanisms to track how those orientations shift over time, may yield more durable and socially responsive outcomes, particularly in addressing the persistent problem of regulatory lag (Wang et al., 2025).

When communicating with the public about AI, presenting contextual details or clarifying system roles alone may not meaningfully engage the public unless those efforts speak to the underlying moral intuitions through which AI risk is processed. Our null results suggest that sensitivity to the specific context of AI use (e.g., healthcare vs. law enforcement) may require deeper moral reflection than public-facing governance communication, such as policy briefs or regulatory proposals, has typically facilitated. One concern is that increasing algorithmic dependence may gradually erode the capacity for such reflection. Vuong and Ho (2025) warn of hidden human costs co-created with AI that accumulate not through dramatic failures but through gradual displacement: as algorithmic systems assume roles previously requiring human judgment, individuals risk progressively losing autonomy over decisions, experiencing degradation of skills that once anchored professional identity, and ceding sacred values that resist quantification to systems optimized for efficiency. Nguyen (2025) raises a related concern through the concept of "order-induced blindness," whereby artificially constructed orders, including those imposed by algorithmic systems, may narrow the interpretive repertoire available to individuals, producing rigid evaluative patterns that resist contextual recalibration. If these dynamics were to take hold, the stable moral templates observed in our data could reflect not careful deliberation but rather a generalized evaluative stance that emerges when the complexity of AI's deployment outpaces individuals' capacity for case-by-case assessment. For governance, this raises a troubling possibility: the very publics whose value frameworks seek to engage may come to operate within constrained evaluative conditions shaped by the technologies under governance. However, this pattern does not imply that contextual reasoning is irrelevant, but rather that existing modes of public communication about AI have not yet succeeded in activating it.

If this interpretation holds, governance strategies must move beyond informational approaches that assume publics will differentiate their responses when given contextual details about AI deployment. Instead, effective governance communication should engage directly with the authenticity concerns and equity intuitions that our findings identify as primary drivers of both concern and regulatory support. In practice, this means designing deliberative processes that create space for moral reflection rather than simply presenting technical risk information, for instance, through citizen assemblies, participatory technology assessments, or structured public consultations that foreground value tradeoffs rather than system specifications. It also suggests that regulatory impact assessments should be expanded beyond technical performance and safety benchmarks to include measures of how proposed AI systems interact with public values such as authenticity, human dignity, and distributional fairness. Without this reorientation, governance risks producing frameworks that are technically sound but normatively illegitimate, satisfying expert criteria while failing to address the concerns that actually shape whether publics accept or resist AI integration into consequential domains of social life.

6.3. Global implications

While this study draws on a U.S. sample, the finding that individual dispositions outweigh contextual and agency cues in shaping AI attitudes carries implications for governance efforts worldwide. If public responses to AI are primarily filtered through deep moral orientations rather than surface-level application cues, then governance frameworks that rely heavily on technical risk classification, however sophisticated, may encounter limitations across national contexts, albeit shaped by different underlying attitudinal structures.

This has direct relevance for international governance initiatives. UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021; updated 2023) explicitly calls for member states to ensure that AI governance reflects cultural diversity and local value systems. Our findings provide empirical support for this principle: if authenticity preferences and equity concerns drive public attitudes more powerfully than application context, then governance frameworks must be calibrated to the dispositional orientations of specific publics rather than assumed to transfer across cultural settings. The Organisation for Economic Co-operation and Development AI Principles (2019, updated 2024) and the World Economic Forum's AI Governance Alliance (2024) similarly emphasize human-centric, inclusive governance, yet our results suggest that such adaptation must extend beyond technical transparency to engage the moral dispositions through which publics actually evaluate AI. For regions currently developing regulatory frameworks, including those in the Global South where AI adoption is accelerating but governance infrastructure remains nascent, these findings suggest that importing risk-classification models wholesale may be insufficient without complementary efforts to understand local value structures. Cross-cultural research on how authenticity, equity, and agency are understood differently across societies is essential for building governance frameworks with genuine democratic legitimacy.

6.4. Limitations and future research

Several limitations should be considered when interpreting these findings. First, the data represent a cross-sectional snapshot collected in March 2025 from a U.S. online panel sample ($N = 1640$). Compared to the most recent U.S. Census data, the sample overrepresents white individuals and skews older than the national median age (U.S. Census Bureau, 2024), and the use of mock online news articles may not fully capture the complexity of real-world message exposure. In terms of measurement, both regulation support and perceived disproportionate impact were assessed using single-item measures. While single-item indicators can lack reliability (Hoeppner et al., 2011), they are generally acceptable when the construct is unidimensional, clearly defined, and narrowly scoped (Fuchs & Diamantopoulos, 2009), as in the present study. Additionally, both dependent variables exhibited relatively high means on 7-point scales, suggesting possible ceiling effects that may

have constrained the range available for experimental manipulations to produce differentiation, while also reinforcing the interpretation that the public has already consolidated strong baseline positions anchored in dispositions. Relatedly, the factorial design may have lacked sufficient statistical power to detect three-way interactions, which typically require substantially larger samples than lower-order effects (Sommet et al., 2023).

The magnitude of agency effects was sensitive to the exclusion of participants who failed the manipulation checks, though dispositional predictors continued to account for substantially more variance under these stricter conditions. This pattern reinforces the central conclusion while suggesting that our estimates of agency effects may be conservative. Full model comparisons are reported in Tables 2 and 3 in the Supplementary Appendix.

These findings point to several directions for future research. First, longitudinal and cross-cultural designs are needed to track how authenticity preferences, equity concerns, and perceptions of agency shift over time and vary across sociocultural settings, particularly given that the dispositions our cross-sectional design treats as stable may themselves be reshaped through sustained interaction with AI-mediated environments. Second, future experimental work should test graded or hybrid forms of human-AI involvement, which may better approximate how people encounter decision-making systems in practice and could reveal context effects that binary agency framings obscure. Third, conceptual work is needed to clarify how authenticity and agency are understood across disciplines, including science communication, human-computer interaction, and moral psychology, as these constructs likely carry different valences in different cultural and intellectual traditions. Finally, given this study's evidence that dispositions outweigh contextual cues, a critical open question is whether structured deliberation can shift this balance. Experimental comparisons of standard informational formats with deliberative interventions would clarify whether contextual reasoning can be activated under conditions that our vignette-based design did not provide, and if so, whether such activation meaningfully redistributes the explanatory weight between dispositional and situational predictors.

7. Conclusion

This study examined whether public attitudes toward AI are shaped more by situational cues, application context, and perceived AI agency, or by stable individual dispositions. Across a factorial experiment with 1,640 U.S. adults spanning medical diagnosis, genetic screening, law enforcement, and screenwriting, neither the context in which AI was deployed nor whether it was framed as autonomous or human-supervised meaningfully shifted public concern or support for regulation. A modest interaction emerged in two contexts: autonomous AI raised concern in genetic screening, while human involvement raised concern in law enforcement, suggesting that agency is interpreted through the cultural meaning of particular applications. Yet these context-specific patterns were small relative to the role of dispositions. Authenticity preference, concern about disproportionate harm to vulnerable groups, and general attitudes toward AI consistently predicted both outcomes across all conditions, indicating that public evaluations of AI in this sample were organized primarily around pre-existing value orientations rather than features of specific applications.

Two implications follow. First, governance approaches that differentiate AI primarily by application context or degree of autonomy may have limited traction with publics whose responses are anchored in dispositions that operate consistently across applications. Engagement strategies that speak directly to authenticity and equity concerns, rather than relying on contextual or technical framing alone, are likely to be more informative for policy design. Second, these dispositions are stable but not static. The cross-sectional design captures them at a single point in time, and sustained interaction with AI-mediated environments may gradually reshape the orientations through which publics evaluate the technology. Findings should therefore be read alongside the study's limitations, including the U.S.-based sample and reliance on vignette-based stimuli. Longitudinal and cross-cultural research is needed to test whether the dispositional patterns observed here persist, attenuate, or vary meaningfully across populations and over time.

CRediT authorship contribution statement

Dayeon Eom: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Julianne Renner:** Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Yijia Erika Zhu:** Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Anqi Shao:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Soobin Choi:** Writing – review & editing, Supervision, Project administration, Data curation, Conceptualization. **Todd Newman:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Dominique Brossard:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Dietram Scheufele:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Ethical approval and informed consent

All procedures were reviewed and approved by the Institutional Review Board of the University of Wisconsin-Madison (2024-1618-CP001).

Funding

Support for this research was provided by the University of Wisconsin–Madison, Office of the Vice Chancellor for Research with funding from the Wisconsin Alumni Research Foundation, and by the Morgridge Institute for Research.

Declaration of conflicting interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.techsoc.2026.103389.

Data availability

Data will be made available on request.

References

Acosta-Enriquez, B. G., Vargas, C. G. A. P., Jordan, O. H., Ballesteros, M. A. A., & Morales, A. E. P. (2024). Exploring attitudes toward ChatGPT among college students: An empirical analysis of cognitive, affective, and behavioral components using path analysis. *Computers and Education: Artificial Intelligence, 7*, Article 100320.

Allyn, B. (2020). *The computer got it wrong': How facial recognition led to false arrest of black man.* NPR. https://www.npr.org/2020/06/24/882683463/the-computer-got-it-wrong-how-facial-recognition-led-to-a-false-arrest-in-michig.

Austin, Z., & Háji, A. (2023). Regulation of wicked problems: Opportunities, responsibilities, and threats. *Journal of Medical Regulation, 109*(3), 6–11.

Bao, L., Krause, N. M., Calice, M. N., Scheufele, D. A., Wirz, C. D., Brossard, D., … Xenos, M. A. (2022). Whose AI? How different publics think about AI and its social impacts. *Computers in Human Behavior, 130*, Article 107182.

Beaudouin, V., Bloch, I., Bounie, D., Clémençon, S., d'Alché-Buc, F., Eagan, J., … Parekh, J. (2020). *Flexible and context-specific AI explainability: A multidisciplinary approach.* arXiv preprint arXiv:2003.07703.

Beets, B., Newman, T. P., Howell, E. L., Bao, L., & Yang, S. (2023). Surveying public perceptions of artificial intelligence in health care in the United States: Systematic review. *Journal of Medical Internet Research, 25*, Article e40337.

Belenguer, L. (2022). AI bias: Exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. *AI and Ethics, 2*(4), 771–787.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press. ISBN 978-1-509-52640-6.

Bui, H. T., Filimonau, V., & Sezerel, H. (2024). AI-thenticity: Exploring the effect of perceived authenticity of AI-generated visual content on tourist patronage intentions. *Journal of Destination Marketing & Management, 34*, Article 100956.

Carman, J., Ballew, M., Lu, D., Leiserowitz, A., Maibach, E., Rosenthal, S., Kotcher, J., Goddard, E., Low, J., Marlon, J., Verner, M., Lee, S., Myers, T., Goldberg, M., Badullovich, N., Mason, T., Aguilar, A., Ongelungel, S. M., Sahlin, K., … Torres, G. (2023). *Climate change in the American mind: Climate justice, spring 2023*. Yale University and George Mason University; Yale Program on Climate Change Communication. https://climatecommunication.yale.edu/publications/climate-change-in-the-american-mind-climate-justice-spring-2023/.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809–825.

Choi, S., Lee, C. J., Park, A., & Lee, J. A. (2025). How the public makes sense of artificial intelligence: The interplay between communication and discrete emotions. *Science Communication, 47*(4), 553–584.

Choung, H., David, P., & Ling, T. W. (2024). Acceptance of AI-powered facial recognition technology in surveillance scenarios: Role of trust, security, and privacy perceptions. *Technology in Society, 79*, Article 102721.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the Center for research and education on aging and technology enhancement (CREATE). *Psychology and Aging, 21*(2), 333–352. https://doi.org/10.1037/0882-7974.21.2.333

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114.

Dorotic, M., Stagno, E., & Warlop, L. (2024). AI on the street: Context-dependent responses to artificial intelligence. *International Journal of Research in Marketing, 41*(1), 113–137.

Dreksler, N., Law, H., Ahn, C., Schiff, D. S., Schiff, K. J., & Peskowitz, Z. (2025). What does the public think about AI? An overview of the public's attitudes towards AI and a resource for future research. *Centre for the Governance of AI*. https://www.governance.ai/research-paper/what-does-the-public-think-about-ai.

Elia, G., & Margherita, A. (2018). Can we solve wicked problems? A conceptual framework and a collective intelligence system to support problem analysis and solution design for complex social issues. *Technological Forecasting and Social Change, 133*, 279–286.

Esmaeilzadeh, P., Mirzaei, T., & Dharanikota, S. (2021). Patients' perceptions toward human–artificial intelligence interaction in health care: Experimental study. *Journal of Medical Internet Research, 23*(11), Article e25856.

European Parliament. (2023). *EU AI Act: First regulation on artificial intelligence*. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

Farooq, U. (2024). Police departments are turning to AI to sift through millions of hours of unreviewed body-cam footage. *ProPublica*. https://www.propublica.org/article/police-body-cameras-video-ai-law-enforcement.

Faverio, M., & Kikuchi, E. (2026). *Key findings about how Americans view artificial intelligence*. Pew Research Center. https://www.pewresearch.org/short-reads/2026/03/12/key-findings-about-how-americans-view-artificial-intelligence/.

Federspiel, F., Mitchell, R., Asokan, A., Umana, C., & McCoy, D. (2023). Threats by artificial intelligence to human health and human existence. *BMJ Global Health, 8*(5).

Ferrer, X., Van Nuenen, T., Such, J. M., Coté, M., & Criado, N. (2021). *Bias and discrimination in AI: A cross-disciplinary perspective. arXiv preprint* arXiv:2008.07309.

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2023). The extent of algorithm aversion in decision-making situations with varying gravity. *PLoS One, 18*(2), Article e0278751.

Fournier-Tombs, E. (2021). Towards a United Nations internal regulation for artificial intelligence. *Big Data & Society, 8*(2), Article 20539517211039493.

Fuchs, C., & Diamantopoulos, A. (2009). Using single-item measures for construct measurement in management research: Conceptual issues and application guidelines. *Betriebswirtschaft, 69*(2), 195–210.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55–130. Academic Press.

Greszki, R., Meyer, M., & Schoen, H. (2015). Exploring the effects of removing "too fast" responses and respondents from web surveys. *Public Opinion Quarterly, 79*(2), 471–503.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Dædalus, 133*(4), 55–66.

Hemesath, S., & Tepe, M. (2024). Multidimensional preference for technology risk regulation: The role of political beliefs, technology attitudes, and national innovation cultures. *Regulation & Governance, 18*(4), 1264–1283.

Hess, D. J. (2015). Publics as threats? Integrating science and technology studies and social movement studies. *Science As Culture, 24*(1), 69–82.

Ho, M. T., & Ho, M. T. (2025). Three tragedies that shape human life in the age of AI and their antidotes. *AI & Society, 40*(7), 5611–5613.

Ho, M. T., Mantello, P., Ghotbi, N., Nguyen, M. H., Nguyen, H. K. T., & Vuong, Q. H. (2022). Rethinking technological acceptance in the age of emotional AI: Surveying Gen Z (Zoomer) attitudes toward non-conscious data collection. *Technology in Society, 70*, Article 102011.

Hoeppner, B. B., Kelly, J. F., Urbanoski, K. A., & Slaymaker, V. (2011). Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. *Journal of Substance Abuse Treatment, 41*(3), 305–312.

Ismatullaev, U. V. U., & Kim, S. H. (2024). Review of the factors affecting acceptance of AI-infused systems. *Human Factors, 66*(1), 126–144.

Janssen, M. (2025). Responsible governance of generative AI: Conceptualizing GenAI as complex adaptive systems. *Policy and Society, 44*(1), 38–51.

Jauernig, J., Uhl, M., & Walkowitz, G. (2022). People prefer moral discretion to algorithms: Algorithm aversion beyond intransparency. *Philosophy & Technology, 35*(1), 2.

Jiang, P., Niu, W., Wang, Q., Yuan, R., & Chen, K. (2024). Understanding users' acceptance of artificial intelligence applications: A literature review. *Behavioral Sciences, 14*(8), 671.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. In *Proceedings of the 28th European conference on information systems (ECIS 2020)*. AIS Electronic Library (Paper 168) https://aisel.aisnet.org/ecis2020_rp/168.

Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society, 9*(1), 20539517221092956.

Kim, S., & Jung, Y. (2023). Development of semantic differential scales for artificial intelligence agents. *International Journal of Social Robotics, 15*(7), 1155–1167.

Kim, J., Yeo, S. K., Brossard, D., Scheufele, D. A., & Xenos, M. A. (2014). Disentangling the influence of value predispositions and risk/benefit perceptions on support for nanotechnology among the American public. *Risk Analysis, 34*(5), 965–980.

Korukonda, A. R. (2005). Personality, individual characteristics, and predisposition to technophobia: Some answers, questions, and points to ponder about. *Information Sciences, 170*(2–4), 309–328.

Lahusen, C., Maggetti, M., & Slavkovik, M. (2024). Trust, trustworthiness and AI governance. *Scientific Reports, 14*(1), Article 20752.

Lee, G., & Kim, H. Y. (2024). Human vs. AI: The battle for authenticity in fashion design and consumer response. *Journal of Retailing and Consumer Services, 77*, Article 103690.

Li, Y., Li, Y., Chen, Q., & Chang, Y. (2024). Humans as teammates: The signal of human–AI teaming enhances consumer acceptance of chatbots. *International Journal of Information Management, 76*, Article 102771.

Loewen, P. J., Lee-Whiting, B., Arai, M., Bergeron, T., Galipeau, T., Gazendam, I., Needham, H., Singer, L., & Yusypovych, S. (2024). *Global public opinion on artificial intelligence (GPO-AI). Schwartz Reisman Institute for Technology and Society, University of Toronto, in collaboration with the Policy, Elections, and Representation Lab (PEARL)*. Munk School of Global Affairs & Public Policy. https://srinstitute.utoronto.ca/public-opinion-ai.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650.

MacMillan, D., Ovalle, D., & Schaffer, A. (2025). Police ignore standards after AI facial recognition matches. *Washington Post*. https://www.washingtonpost.com/business/interactive/2025/police-artificial-intelligence-facial-recognition/.

Maese, E. (2025). Americans use AI in everyday products without realizing it. Gallup. https://news.gallup.com/poll/654905/americans-everyday-products-without-realizing.aspx.

Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change, 175*, Article 121390.

McClain, C., Kennedy, B., Gottfried, J., Anderson, M., & Pasquini, G. (2025). *How the U.S. public and AI experts view artificial intelligence*. Pew Research Center. https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/.

Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., … Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health, 2*(6), Article e0000278.

Newman, G. E. (2019). The psychology of authenticity. *Review of General Psychology, 23*(1), 8–18.

Nguyen, M.-H. (2025). Daoist thoughts and Kingfisherish musings. *SSRN*. https://doi.org/10.2139/ssrn.5975337 [Working paper].

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867–872.

Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & Society, 35*(4), 927–936.

Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. *Nature Medicine, 30*(11), 3098–3100.

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes nordic national public policy strategies for artificial intelligence (AI). *Technology in Society, 63*, Article 101421.

Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. In *Companion proceedings of the 2019 world wide web conference* (pp. 539–544).

Sartori, L., & Bocca, G. (2023). Minding the gap(s): Public perceptions of AI and socio-technical imaginaries. *AI & Society, 38*(2), 443–458.

Schepman, A., & Rodway, P. (2020). Initial validation of the general attitudes towards artificial intelligence scale. *Computers in Human Behavior Reports, 1*, Article 100014.

Seyama, J. I., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence, 16*(4), 337–351.

Sommet, N., Weissman, D. L., Cheutin, N., & Elliot, A. J. (2023). How many participants do I need to test an interaction? Conducting an appropriate power analysis and

achieving sufficient power to detect an interaction. *Advances in Methods and Practices in Psychological Science, 6*(3), 25152459231178728.
Strauß, S. (2021). Deep automation bias: How to tackle a wicked problem of AI? *Big Data and Cognitive Computing, 5*(2), 18.
Stückelberger, C. (2024). Balanced global and contextual ethical values: A comparison of official national value systems and proposals for the new context. *Vietnamese Journal of Philosophy, 3*(69), 47–57.
Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human-AI interaction (HAII). *Journal of Computer-Mediated Communication, 25*(1), 74–88.
Sundar, S. S., & Kim, J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–9). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300768. Paper No. 538.
Sweet, W. (2023). National values, culture, and philosophy. *Philosophy, 65*(1), 55–64.
Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society, 40*(2), 137–157.
Tyson, A., & Kikuchi, E. (2023). *Growing public concern about the role of artificial intelligence in daily life*. Pew Research Center. https://www.pewresearch.org/short-reads/2023/08/28/growing-public-concern-about-the-role-of-artificial-intelligence-in-daily-life/.
Tyson, A., Pasquini, G., Spencer, A., & Funk, C. (2023). *60% of Americans would be uncomfortable with provider relying on AI in their own health care*. Pew Research Center. https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/.
U.S. Census Bureau. (2024). *Annual estimates of the resident population by sex, age, race, and Hispanic origin for the United States: April 1, 2020 to July 1, 2023 [data set]*. U.S. Department of Commerce. Retrieved May 5, 2026, from https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html.
Valčo, M. (2024). Impact of social media platforms' AI algorithms on human values. *Vietnamese Journal of Philosophy, 4*(70), 53–64.
Van Berkel, N., Tag, B., Goncalves, J., & Hosio, S. (2022). Human-centred artificial intelligence: A contextual morality perspective. *Behaviour & Information Technology, 41*(3), 502–518.
Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, 1–67.
Verma, P. (2022). The never-ending quest to predict crime using AI. *Washington Post*. https://www.washingtonpost.com/technology/2022/07/15/predictive-policing-algorithms-fail/.
Vo, D. T., Nguyen, L. V. T., Dang-Pham, D., & Hoang, A. P. (2025). What makes an app authentic? Determining antecedents of perceived authenticity in an AI-powered service app. *Computers in Human Behavior, 163*, Article 108495.
Vuong, Q. H., & Ho, M. T. (2025). Abundance of words versus poverty of mind: The hidden human costs co-created with LLMs. *AI & Society, 40*(2), 1137–1138.
Vuong, Q. H., La, V. P., & Nguyen, M. H. (2025). Informational entropy-based value formation: A new paradigm for a deeper understanding of value. *Evaluation Review, 50*(3), Article 0193841X251396210.
Vuong, Q. H., & Napier, N. K. (2015). Acculturation and global mindsponge: An emerging market perspective. *International Journal of Intercultural Relations, 49*, 354–367.
Wang, X., Xie, F., & Liu, B. (2025). Governance efficiency and upgrade pathways of international generative AI policies and regulations. *Technology in Society*, Article 103082.
Wilson, C., & Van Der Velden, M. (2022). Sustainable AI: An integrated model to guide public sector decision-making. *Technology in Society, 68*, Article 101926.
World Economic Forum. (2024). AI governance alliance: Briefing paper series. https://www.weforum.org/publications/ai-governance-alliance-briefing-paper-series/.
Wu, J., Xu, L., Yu, F., & Peng, K. (2021). Acceptance of medical treatment regimens provided by AI vs. human. *Applied Sciences, 12*(1), 110.
Yan, W., Hu, B., Liu, Y. L., Li, C., & Song, C. (2024). Does usage scenario matter? Investigating user perceptions, attitude and support for policies towards ChatGPT. *Information Processing & Management, 61*(6), Article 103867.
Yang, J., & Battocchio, A. F. (2021). Effects of transparent brand communication on perceived brand authenticity and consumer responses. *The Journal of Product and Brand Management, 30*(8), 1176–1193.
Yue, B., & Li, H. (2023). The impact of human-AI collaboration types on consumer evaluation and usage intention: A perspective of responsibility attribution. *Frontiers in Psychology, 14*, Article 1277861.
Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*. Available at: SSRN 3312874.