Journal of Language, Media and Society
Vol. 1, No. 1 https://doi.org/10.66056/jlms.2026.v1i1.002

Carter & Co.
PUBLISHING

WHEN SYSTEMS MISRECOGNIZE THEIR USERS: A SEMI-SYSTEMATIC REVIEW OF COMMUNICATION, IDENTITY, AND BIAS IN LLMS

LIJING GAO & RUANJIA LIU

Large language models (LLMs) offer vast computational power yet consistently overlook cultural nuance. While often praised for bridging language gaps, recent research highlights a deeper issue: LLMs largely reflect Anglophone and Western European cultural values, which are embedded in the English-dominated data that shapes them. This review synthesizes findings from 2020 to 2025 to assess the implications of this technological spread for vulnerable groups, particularly immigrants, refugees, and international students, who must navigate adaptation within English-centric and Western communication norms.

Using insights from cultural cognition and identity-protective cognition theories, the analysis identifies five central forms of bias: (1) representational bias that undermines non-Western perspectives; (2) linguistic inequity that amplifies challenges for low-resource languages; (3) authenticity failures, with stereotypes substituting for real cultural understanding; (4) identity erosion as users' voices are homogenized; and (5) reliance on LLMs that may hinder independent language skill development. This "equity paradox" means that the very systems marketed as democratizing global communication can actually deepen exclusion and sameness among those who are most reliant on them.

Ultimately, the review concludes that current governance and policy efforts are insufficient to address the underlying power dynamics that shape LLM development. Authentic cross-cultural communication, the evidence suggests, depends on human qualities absent in LLMs: presence, vulnerability, and the openness to change that underpins accurate understanding. In an AI-mediated world, recognizing the limits of these tools is not a matter of nostalgia, but rather necessary wisdom.

KEYWORDS: large language models; cross-cultural communication; cultural bias; identity formation; linguistic diversity; cultural homogenization; digital equity; AI governance; cultural representation

Introduction: The Seduction and Betrayal of Technological Solutions

Current discussions regarding large language models (LLMs) frequently resemble a form of technological idealism, as innovative computer tools pledge to eliminate the obstacles that have historically hindered human interaction across linguistic and cultural divides. The narrative is compelling: LLMs such as ChatGPT and Claude serve serve as democratic instruments, accessible worldwide, enabling nonnative speakers to communicate with unparalleled fluency, tapping into cultural insights and linguistic refinement that were once attainable solely through extensive immersive experience (Hu et al., 2025; Khan et al., 2025; Li et al., n.d.; Lo et al., 2024; Mao et al., 2025; Zangana et al., 2025; Zhang et al., 2025; Zohouri et al., 2024). From this perspective, LLMs appear almost redemptive, promising to correct historical inequities in global communication systems.

Yet this narrative conceals more troubling realities. Behind the outer layer of acceptable performance metrics lies an intrinsic bias that privileges certain viewpoints, assessments, and communication styles while marginalizing others (Ashraf et al., 2025; Keleg, 2025; Shan et al., 2024). The algorithms being implemented worldwide were predominantly trained on English-language materials authored mainly by Anglophone writers, reflecting Anglophone concerns. When asked to respond as an "average person," GPT models align most closely with cultural values from Finland, Andorra, and the Netherlands while showing the greatest distance from Jordan, Libya, and Ghana (Dairo, 2005; Ferdaus et al., 2024; Liu et al., 2025; Lyu & Du, 2025). These are not random variations. They represent systematic underrepresentation of non-Western moral frameworks, epistemologies, and ways of understanding the world.

For newcomers navigating life in a foreign country, whether immigrants working to understand local social norms, refugees grappling with cultural adaptation, or international students seeking connection in unfamiliar settings, these technologies present a particularly intricate set of challenges. On one side, they provide authentic support: instruments that aid non-native speakers in generating more fluent English, alleviating the cognitive load of translation, and expediting language acquisition (Ehrensberger-Dow et al., 2020; Elgamal, 2019). Conversely, these instruments may consistently misrepresent local culture, perpetuate detrimental preconceptions, standardize unique expressions, and foster dependencies that compromise the independent proficiency necessary for enduring cultural adaptation (Algouzi & Alzubi, 2023; Sahebi & Formosa, 2025).

This article presents a semi-systematic review of the dual promise and peril of large language models in cross-cultural contexts. The review synthesizes emerging evidence showing that cultural bias is not a peripheral flaw but an intrinsic feature of these systems to shaped by their training data, parameter settings, optimization processes, and institutional uses. Across the literature, a consistent pattern emerges regarding who benefits from current deployment practices and who disproportionately bears the costs. At its core, this review asks what genuine equity would require in AI-mediated communication.

Semi-Systematic Review Rationale and Methods

Rationale

This review employs a semi-systematic methodology, which combines the transparency and reproducibility of systematic review protocols with the methodological flexibility necessary for research in rapidly evolving domains (Snyder, 2019). Unlike fully systematic reviews that require exhaustive search strategies and rigid adherence to all PRISMA guidelines (Page et al., 2021), semi-systematic approaches allow targeted inclusion criteria and pragmatic methodological adjustments, particularly well-suited to interdisciplinary and fast-moving fields such as large language model research (Snyder, 2019). This approach balances the need for evidence synthesis rigor with recognition that emerging research areas may benefit from focused rather than exhaustive searches.

Inclusion Criteria

Studies were eligible for inclusion referred to the following pre-specified criteria, established following PRISMA 2020 guidelines (Page et al., 2021):

Publication timeframe: Studies published between January 2020 and October 2025 were included. This five-year window captures the post-ChatGPT era beginning with its November 30, 2022 release (OpenAI, 2022), which catalyzed unprecedented research interest in generative artificial intelligence. All literature searches were conducted across multiple databases, including Google Scholar, Web of Science, and arXiv. While Web of Science and arXiv returned relevant results, Google Scholar provided the most comprehensive coverage by aggregating papers from diverse sources relevant to this research. A five-year search window captures rapidly evolving fields through major technological transitions while remaining operationally feasible. (Hamman et al, 2017).

Study design: Empirical research was prioritized, including studies examining LLM use in cross-cultural or multilingual contexts as well as broader research on cross-cultural communication and cultural adaptation challenges that provide contextual framework for understanding LLM applications. Preprints and conference proceedings were included to ensure a comprehensive and timely evidence synthesis in this rapidly evolving field. In fast-moving technological domains, traditional peer-review timelines can render findings outdated before publication (Suh, 2025). The COVID-19 pandemic established clear precedent for incorporating preprints into systematic reviews, showing that preprints exhibit minimal discrepancies from their later published versions and can improve estimate precision while enabling timely dissemination (Tennant et al., 2018). In AI research in particular, platforms such as arXiv now serve as central communication channels (Suh et al., 2025), and major institutions and technology companies routinely use preprints to accelerate knowledge transfer (Suh et al., 2025; Sætra, 2024). This practice aligns with open science principles that prioritize transparency, accessibility, and rapid dissemination (Tong et al., 2025; UNESCO, 2021), which are particularly critical for evidence synthesis addressing contemporary AI applications where policy and practice decisions require current information (Yang et al., 2025). Conference proceedings were included because they provide access to emerging research and diverse perspectives that may not yet appear in journal publications, helping reduce publication bias and identify relevant evidence comprehensively (Scherer & Saldanha, 2019). Empirical evidence to defined as research based on observed and measured phenomena rather than theory or belief (Cook et al., 1997) to provides the strongest foundation for evidence synthesis and supports valid conclusions regarding real-world applications (Gopalakrishnan & Ganeshkumar, 2013; Tong et al., 2025). This inclusive approach allows integration of relevant contextual research that clarifies the cultural and communicative dimensions within which LLMs operate, while maintaining methodological rigor through systematic screening and quality assessment procedures.

Language: English-language studies were included due to resource constraints. This pragmatic inclusion criterion reflects the semi-systematic approach, which balances rigor with operational feasibility (Snyder, 2019; Pham et al., 2005), though the resulting language bias is acknowledged as a study limitation.

Exclusion Criteria

Studies were systematically excluded if they: (1) lacked cultural analysis (purely technical or algorithmic papers without examination of cultural, linguistic, or social dimensions); (2) were non-empirical (opinion pieces, commentaries, editorials, or theoretical work without original data); (3) examined exclusively monolingual contexts as the review focuses specifically on cross-cultural and multilingual dimensions of LLM use and cultural adaptation unless they provided empirical evidence on communication or cultural adaptation challenges directly applicable to understanding LLM performance in cross-cultural contexts. As a result, this research identified 536 articles into analysis.

The methodological workflow for this semi-systematic review is depicted in Figure 1, highlighting the pragmatic integration of PRISMA guidelines with targeted inclusion strategies to synthesize empirical research from January 2020 to October 2025.

Figure 1: Semi-Systematic Review Methodology and Study Selection Flow

Phase	Action & Criteria Details
1. Identification	**Database Search (Jan 2020 – Oct 2025)** Sources: Google Scholar (Primary), Web of Science, arXiv. Focus: Post-ChatGPT era & technological transitions.
2. Screening	**Inclusion Criteria (Pragmatic & Targeted)** **Type:** Empirical research, preprints, & conference proceedings. **Scope:** LLMs in cross-cultural/multilingual contexts; cultural adaptation. **Language:** English only (operational feasibility).
3. Eligibility	**Exclusion Criteria** 1. **Technical Only:** Lacks cultural/social analysis. 2. **Non-Empirical:** Editorials, opinions, or purely theoretical work. 3. **Monolingual:** Excludes studies without cross-cultural dimensions.
4. Included	**Final Sample Size: N = 536 Articles** Synthesized to address three core Research Questions (RQs).

This semi-systematic review examines how large language models mediate cross-cultural communication and in which contexts their limitations become most consequential. Drawing on the evidence gathered, the review is structured around three central research questions: **(1)** How do LLMs facilitate or hinder cross-cultural and multilingual communication for users navigating unfamiliar cultural contexts? **(2)** What cultural biases and worldview misalignments are systematically embedded in LLM outputs, and how do these patterns shape users' trust, interpretation, and reliance on these systems? **(3)** What would genuine communicative equity require in the design, training, and governance of LLMs used across diverse cultural settings? These questions organize the synthesis that follows and provide a framework for interpreting patterns across the empirical literature.

Results

Initial analysis of the corpus confirms that the functional benefits of LLMs for cross-cultural communication are substantial and widely documented. For users operating outside their dominant language, these systems act as critical cognitive scaffolds, significantly reducing the mental burden of translation and syntactic formulation (Lee et al., 2024). By elevating lexical precision and smoothing linguistic roughness, LLMs effectively democratize access to high-stakes professional and academic discourse, offering a provisional form of linguistic equity that allows newcomers to bypass traditional gatekeeping mechanisms (Nguyen et al., 2024; Liang et al., 2023). However, the literature increasingly suggests that this improved surface-level accessibility frequently comes at the cost of deeper semantic integrity.

Section I: The Architecture of Bias—How Cultural Representation Becomes Cultural Erasure

1.1 The Manifestation of Western-Centrism

One might begin with a simple question: What does a language model "believe"? The answer reveals itself not in explicit statements but in consistent patterns of preference and omission. Systematic evaluation across diverse LLM architectures reveals a highly consistent pattern across the majority of studies: when asked to engage with cultural questions, these systems tend to prioritize values characteristic of Protestant European and Anglo-American contexts (Segerer, 2025).They default to self-expression values such as environmental protection, tolerance of diversity, and gender equality, which predominate in wealthy Western societies, even when prompted in non-Western languages or explicitly instructed to adopt non-Western perspectives. In what follows, we use a series of illustrative, composite scenarios to concretize patterns observed across the empirical literature. These examples are not single case reports but theoretically and empirically informed vignettes that synthesize recurring dynamics identified in prior studies.

This is not neutral linguistic performance. It is cultural transmission. When LLMs repeatedly associate Indian women with domestic labor, Russians with vodka consumption, or Arabs with terrorism, they do not merely reflect societal biases that happened to appear in training data (Wu et al., 2025). Rather, they crystallize and scale those biases across millions of interactions, routinizing and legitimizing them through the appearance of technological objectivity.

Consider the consequences for newcomers. Unfamiliar with local social conventions such as professional clothing, conversational indirectness, and friendship duties, an immigrant confronts an LLM schooled mostly in Western models and frameworks. When she asks about dinner etiquette or gift-giving customs, she receives advice filtered through Western preoccupations and Western categorical systems (Pedersen et al., 2025). The system has little understanding of her destination culture because few literature in that language describe it well enough to modify model parameters. Instead of silence or doubt, she receives information that appears confidence, cultural expertise, and neutral information. Only after social failures and misunderstandings does she realize the advice was misleading.

The mechanisms that contribute to the social cohesion bias are well understood. Large language models are trained on textual data, which is significantly unbalanced within global information systems. English prevails, despite accounting for approximately 15% of the global population (Crystal, 2003; Rao, 2019). Western perspectives prevail, even though they account for a diminutive fraction of human diversity. The training data reflects this imbalance, not due to intentional decisions but as a result of the inherent structures of digital information distribution. Once established, this imbalance proves to be notably challenging to remove. Models retain these biases, even when trained on multilingual data (Ashraf et al., 2025; Keleg, 2025). Despite explicit instructions to incorporate non-Western perspectives, they process requests through frameworks influenced by their training. The bias is an inherent characteristic of the architecture rather than a correctable error.

1.2 The Hidden Mechanisms: How Bias Persists Despite Efforts to Address It

Understanding cultural bias in large language models requires examining mechanisms operating at multiple levels, often invisible to users and insufficiently addressed by current interventions. This review traces the process through which bias is generated and perpetuated to from imbalances in training corpora and tokenization, through model objectives and alignment procedures, to interface design and institutional deployment practices that normalize some voices while marginalizing others.

Data Composition and Representational Asymmetry

Across the studies reviewed, there is broad and consistent evidence that the foundational source of cultural bias in large language models stems from a profound asymmetry in their training corpora. Multiple analyses show that these datasets are dominated by English-language texts drawn from Western media, publishing, and digital platforms (Ferdaus et al., 2024; Ghimire, 2025; Han et al., 2025; Liu et al., 2025; Lyu & Du, 2025). The imbalance is striking: although English accounts for roughly 15 percent of global languages, it makes up an estimated 50 to 70 percent of the digitized text used to train major LLMs (Segerer, 2025; Lehdonvirta, 2022; Chatterji et al. 2025). In contrast, low-resource languages, typically those with fewer than ten million speakers and limited digital documentation, contribute less than five percent of most training datasets, even though they represent the home languages of billions of people (Ferdaus et al., 2024; Liu et al., 2025; Lyu & Du, 2025).

The problem extends beyond quantitative imbalance to what might be described as a representational hierarchy. The texts included in training corpora reflect not only the distribution of languages but also the knowledge systems that are recognized as authoritative. Academic publishing, news media, and digitized books, which serve as the primary sources of training data, are concentrated in Western institutions and largely encode Western epistemological frameworks. Indigenous knowledge systems, non-Western philosophical traditions, and alternative ways of understanding fundamental concepts such as causality, time, personhood, and ethics are either absent from training data or appear only when mediated through Western interpretive frameworks and scholarly representations (Abdilla & Crawford, 2020; Birhane et al., 2022; De Sousa Santos, 2014; Gwagwa et al., 2020; Hoppers, 2002; Kamran, 2024; Leibo et al., 2025; Peters & Carman, 2024). For example, a Navajo environmental practice may be recorded not by Navajo knowledge holders but by Western environmental scientists describing it; the model therefore learns the Western account rather than the Indigenous understanding. This compositional bias means that the model is trained not on the full diversity of global human knowledge but on a particular, narrowly Western subset that is presented as universal.

When large language models are trained on these compositionally biased datasets, the statistical regularities within the data become encoded not as explicit rules or retrievable statements but as learned parameters distributed across millions of artificial neurons. This marks a critical point of translation: the cultural preferences embedded in the data are mathematically instantiated in the model's functional architecture. Consider concretely what this means. When training data disproportionately associates certain concepts with certain cultural contexts, for example linking leadership with masculine pronouns and Western business terminology (Garg et al., 2018; Müller et al., 2025), associating subsistence practices with non-Western peoples (Malu, 2025), or repeatedly pairing development with Western-style industrialization (Malu, 2025), these associations accumulate into learned representations. The model develops what might be called vector space preferences: patterns in how concepts relate to one another in high-dimensional geometric space (Schröder et al., 2024; Müller et al., 2025). Large language models are 3-6 times more likely to recommend occupations that stereotypically align with a person's perceived gender, with boys receiving substantially more STEM-related career suggestions than girls (Torres et al., 2023; Fock & Siller, 2025). These patterns become part of how the model understands the world, functioning not as programmed bias but as learned associations that disadvantage marginalized groups through statistical co-occurrence patterns in training corpora (Bender et al., 2025; Torres et al., 2023).

The particular challenge with this encoding is that these learned parameters are not easily inspected, modified, or removed. A programmer cannot simply edit individual parameters the way they might edit explicit rules in traditional software. The biases are emergent properties of millions of parameters working together, making them what scholars call implicit knowledge to knowledge that shapes behavior but isn't easily articulated or adjusted. When researchers attempt to debias models after training, they face the problem that the bias isn't localized to one parameter or one layer but distributed throughout the network. Attempts to reduce specific stereotypes through post-training techniques often fail because the underlying statistical patterns remain embedded in the model's fundamental structure (Glickman & Sharot, 2025). To meaningfully address bias at the parameter level often requires architectural redesign or retraining, which are computationally expensive and practically difficult to implement at scale.

Prompt conditioning activates and adjusts these priors without altering their essential nature. Users trying to direct models towards more culturally sensitive outputs through rigorous prompting face a limitation: the model can modify emphasis and nuance, but cannot beyond the inherent viewpoints ingrained in its training (Agarwal et al. 2025; Liang et al; 2025; Shen et al., 2025). An Arabic speaker prompting in Arabic continues to receive outputs that exhibit Western biases, as the model's depiction of Arabic culture has been influenced by English-language texts regarding Arab culture, which are filtered through Western frameworks and frequently perpetuate Orientalist stereotypes (Alyafeai et al., 2023; Sallam & Mousa, 2024).

Human feedback loops often amplify the problem rather than resolve it. When companies apply reinforcement learning from human feedback, a process in which human evaluators rate model outputs to guide further training, a critical question emerges: who are these humans? The research is clear: they are disproportionately Western, educated, English-speaking (Lodoen & Orchard, 2025). Their preferences become embedded in the model's objectives. The system learns not to produce correct or culturally appropriate responses, but to produce responses that align with the preferences of these specific human evaluators. Bias becomes recursively reinforced (Wang et al., 2024; Glickman & Sharot, 2025).

Downstream adoption institutionalizes and scales the bias. When universities, employers, and government agencies deploy these systems, they do not simply use a tool to they enact and perpetuate the model's embedded cultural assumptions (Bao et al. 2025; Prakash et al. 2025; Zheng, 2024; Garcia, 2025). Students learn using systems that treat Western ways of thinking as standard and other ways as exotic. Employees write using tools that systematically alter their distinctive voices toward Western norms. Citizens seek government information from systems trained on Western legal and administrative frameworks.

The result is an escalating process wherein cultural bias, far from being a marginal concern, sits at the very heart of how these systems operate. Addressing it requires not tinkering at margins but fundamental rethinking of how models are trained, evaluated, and deployed.

Section II: The Lived Experience—What Bias Means for Newcomers

2.1 Cultural Misrepresentation and the Loss of Meaning

A growing body of research demonstrates that LLMs routinely reshape nonnative speakers' writing in ways that privilege Western academic styles and epistemic priorities. The scenario of an Iranian graduate student refining her explanation through ChatGPT reflects a pattern observed across multiple empirical evaluations: while the language becomes more fluent, the system often redirects emphasis, reframes arguments, or omits culturally grounded reasoning. The resulting text is polished yet noticeably aligned with Western communicative conventions, illustrating a phenomenon repeatedly documented in cross-cultural studies of LLM use.

This entails cultural misrepresentation, albeit through subtle, nearly imperceptible modifications. Pedersen and colleagues' research demonstrated that LLMs consistently struggle with the interpretation of culture-specific metaphors and idioms (Pedersen et al., 2025). When prompted to elucidate a Danish phrase grounded in Danish history and culture, both ChatGPT and Llama encountered difficulties, either wrongly incorporating English metaphors or resorting to ambiguous generalizations. The sentiment becomes "lost in translation," not due to linguistic disparities but because the model lacks the historical and cultural acumen required to comprehend how that word encapsulates an entire worldview.

For newcomers, the consequences are compounded. An immigrant trying to understand what it means to "work hard" in a new culture does not encounter the nuanced, context-specific meanings that exist in reality, but instead a simplified, Westernized narrative, which centered on the Horatio Alger story and the mythology of pulling oneself up by the bootstraps. This narrative contrasts sharply with how work and effort are understood in many non-Western cultures, where collective responsibility, family honor, or spiritual purpose may take precedence. Yet the large language model presents its version as a neutral and factual account.

The failure extends beyond metaphor. Multilingual proficiency does not ensure cultural representation: research shows AI systems marginalize African cultural expressions through Eurocentric training data (Bignotti, 2025), Latin American media lack localized AI imagery despite multilingual coverage (Sanguinetti & Palomo, 2025), and meaningful representation requires community-grounded evaluation rather than technical capability alone (Qadri et al., 2025). U.S.-centric bias persisted even when models were prompted in their native languages. The researchers demonstrated that self-consistency was a stronger predictor of intercultural alignment than multilingual competence alone, suggesting that the problem lies not in linguistic translation but in the deeper cultural frameworks guiding how the model processes information.

2.2 Stereotype Reinforcement and the Crystallization of Prejudice

Across the studies included in this review, there is substantial and converging evidence that LLMs reinforce culturally embedded stereotypes in ways that meaningfully influence downstream behavior. Pareek (2025) assessed prominent LLM systems with psychology-based metrics particularly formulated to identify biases. Their findings indicate that even systems intentionally meant to be "value-aligned" displayed systematic preconceptions related to race, gender, religion, health, and other variables.

The troubling part: these word association biases proved diagnostic of downstream discriminatory behavior. Models that showed stereotype bias in controlled tests also generated more biased, inappropriate, or harmful outputs in real-world applications. Stereotypes do not exist harmlessly in latent space. They shape what the system produces when deployed.

For a refugee from Myanmar seeking to understand his place in his new community, these stereotypes are not abstract concerns. When the LLM he consults for advice about workplace relationships consistently associates his background with certain traits or capabilities, whether explicitly or through subtle framing, those biases shape both what information he receives and how he begins to perceive himself. Research on stereotype threat reveals that these impacts are not purely individual; they accumulate through repeated interactions with systems that encode and perpetuate prejudice (Wang et al., 2024; Vasista et al., 2025).

Researchers documented an even more troubling phenomenon: stereotypes can arise spontaneously during LLM-based multi-agent interactions, even when the individual agents begin without any predefined bias. The strength of these stereotypes increases within hierarchical systems and through repeated exchanges (Guo & Xu, 2025; Mehdizadeh & Hilbert, 2025), following the same dimensions of warmth and competence that appear consistently across architectures such as GPT, Claude, Mistral, DeepSeek, and Gemini (Guo & Xu, 2025; Borah & Mihalcea, 2024). These findings suggest that stereotype formation is not a model-specific artifact but a structural feature of how large language models learn from and interact with one another (Haase & Pokutta, 2025; Binkyte, 2025). Addressing these emergent biases requires value alignment frameworks that account for multi-agent dynamics rather than focusing solely on individual model behavior (Zeng et al., 2025).

2.3 Communication Style Mismatch and the Imposition of Inappropriate Norms

Although studies vary in their methodological approaches, a clear pattern emerges across the literature: LLMs often misalign with the communication norms of cultures that rely on indirectness, relational cues, or high-context signaling. Havaldar and colleagues' Culturally-Aware Conversations framework identifies three dimensions that shape communication: situation, relationship, and cultural background (Havaldar et al., 2025).

Large language models perform well in direct, low-context cultures such as the United States and the Netherlands but struggle in societies where indirectness signals respect, humility outweighs self-promotion, and formality conveys attentiveness. Trained mainly on Western norms, they promote behaviors that conflict with local expectations.

Section III: The Equity Paradox—When Tools for Empowerment Become Mechanisms of Exclusion

3.1 The Cruel Irony: When Non-Native Speakers Must Use AI to Avoid AI Accusations

Recent research revealed a troubling finding: AI detection methods incorrectly label 61.3% of essays authored by non-native English speakers as AI-generated (Liang et al., 2023; Jiang et al., 2024). The mechanism is straightforward: non-native authors inherently exhibit reduced linguistic diversity, diminished syntactic complexity, and decreased lexical richness compared to native speakers to the same qualities characteristic of AI-generated writing (Fraser et al., 2025; Lege, 2024). Detection systems, unable to discriminate between sources of reduced complexity, classify human effort as artificial.

The cruelty of this situation is clear: to avoid false accusations of AI use, non-native writers are compelled to use AI to augment their linguistic diversity. To be recognized as genuinely human, they must first become augmented by AI. This technical paradox threatens the inclusion of non-native English speakers in global academic and professional spheres precisely when access becomes most critical (Jiang et al., 2024; Lege, 2024).

Liang and colleagues quantified this paradox: when ChatGPT improved TOEFL essays to resemble native-speaker writing, the average false positive rate dropped by 49.7%, from 61.3% to 11.6% (Liang et al., 2023). The implication is clear: linguistic diversity functions as a proxy for human authenticity in detection systems, yet non-native speakers inherently lack that diversity due to their ongoing development of linguistic competence (Fraser et al., 2025).

3.2 The Double Jeopardy: Cost Barriers and Performance Degradation

Across diverse technical and empirical studies, there is strong evidence that LLMs systematically disadvantage low-resource languages. Speakers of low-resource languages incur costs that are 4-6 times higher per unit of usage compared to English speakers (Solatorio et al., 2024). This situation illustrates the mechanics of tokenization: languages underrepresented in training data are divided into numerous separate tokens, each consuming additional computational resources (Ahia et al., 2023; Petrov et al., 2023). The result is higher costs paired with diminished returns.

Concurrently, the performance of LLMs significantly diminishes for low-resource languages (Petrov et al., 2023; Rahman et al., 2024). Translation quality metrics reveal substantial performance declines in low-resource languages (Solatorio et al., 2024; Pakray et al., 2025). Healthcare information obtained from LLMs in non-English languages demonstrates significantly lower accuracy and utility: correctness decreases by 18% in Spanish, Chinese, and Hindi compared to English, with non-English responses 29% less consistent than their English counterparts (Chandra & Jin, 2024). This creates what researchers term double jeopardy to individuals with the fewest resources face the greatest obstacles and poorest outcomes, fundamentally undermining equitable technology access (Solatorio et al., 2024).

3.3 Detection Bias and Response Quality Disparities

Across the literature reviewed, multiple independent studies provide compelling evidence that LLM outputs shift systematically in response to linguistic identity cues and cultural context. A Nature study using matched-guise prompts found that models make systematically more negative decisions for text written in African American English, despite no overt mention of race (Hofmann et al., 2024). A cross-country audit in *PNAS Nexus* showed that default outputs cluster toward English-speaking/Protestant European values, with cultural prompting only partly reducing this bias (Tao et al., 2024). Randomized trials also document anchoring effects in LLMs and show that outputs move when primed by preceding information, underscoring how seemingly minor cues can alter model quality and judgments (Nguyen, 2024). Together, these results indicate that disparities reflect dialect and cultural alignment rather than merely "nativeness," and can be triggered by simple primes.

3.4 The Workplace Dynamics: When Quality Degradation Becomes Social Stigma

Although the published literature on workslop, meaning superficially polished but substantively weak AI-generated workplace content, is still developing, the available evidence shows a consistent and widely recognized pattern of social and organizational consequences. Early studies report that AI-mediated, low-quality responses erode trust among coworkers, requiring recipients to verify accuracy, rewrite vague sections, and navigate the interpersonal discomfort of questioning a colleague's contribution (Hancock & Niederhoffer, 2025). Broader organizational research reinforces these concerns: approximately 40% of workers report rising workloads associated with such breakdowns (Niederhoffer et al., 2025; Richardson & Antonello, 2022), which reliably undermines trust and reduces willingness to collaborate. These burdens fall disproportionately on non-native speakers, who rely on LLMs for linguistic fluency (Brynjolfsson et al., 2025) but whose use of AI tools may be misinterpreted as lack of effort or cultural awareness (Nguyen et al., 2024). Taken together, the emerging evidence indicates that technologies intended to support communication can inadvertently generate stigma and exacerbate inequities (Niederhoffer et al., 2025; Koo, 2025).

3.5 Educational Performance Gaps and Linguistic Brittleness

Across the studies examined, there is **substantial and recurring evidence** that equity concerns emerge in educational settings where students rely on large language models as learning tools. While some models demonstrate relatively stable performance across multiple languages, numerous evaluations report a marked drop in accuracy and instructional quality when the models operate in underrepresented or low-resource languages. Kwak and Pardos (2024), for example, document consistent underperformance for languages such as Irish and Marathi compared to English-based educational taxonomies, reflecting a broader pattern identified across the literature.

Rodrigues and colleagues evaluated LLM performance in answering educational questions in Brazilian Portuguese across different question types, subjects, and difficulty levels (Rodrigues et al., 2025). Their preliminary findings suggest potential for LLMs to support diverse educational needs, though performance varies by question characteristics. Separate research has demonstrated that LLM performance in English consistently surpasses that of other languages (Solatorio et al., 2024; Ahia et al., 2023), and that employing prompts in English often produces better outcomes even in non-English contexts, indicating that LLM representations are primarily shaped by English-centric training frameworks (Tao et al., 2024).

Recent research reveals significant concerns about LLM brittleness: model performance varies dramatically (±23%) across common benchmarks when only single characters such as delimiters between examples are modified, despite semantic information remaining identical (Su et al., 2025). This brittleness is not unique to minute formatting changes; LLMs show unpredictable performance across prompt phrasing, semantic structure, and element ordering on even elementary tasks like set membership queries (Hergert et al., 2025). While humans exhibit some sensitivity to prompt modifications, LLMs demonstrate greater instability, particularly to typographical errors and label order reversals (Li et al., 2025).

Current benchmarks emphasize standardized, formal writing that inadequately reflects diverse human communication styles, resulting in limited external validity for real-world performance assessment. This brittleness has profound consequences for immigrant populations, international students, and multilingual professionals: AI-mediated guidance and evaluation systems affect academic achievement, job opportunities, and social integration when applied across varied linguistic and stylistic contexts (Su et al., 2025; Hergert et al., 2025).

Section IV: Identity Erasure—How Technology Homogenizes Culture

4.1 The Systematic Stripping of Identity Markers

Language functions simultaneously as a communicative instrument and a primary marker of cultural affiliation and self-perception, profoundly shaping how immigrants negotiate identity in new sociocultural environments (Darginavičienė, 2023; Fielding, 2021; Gao, 2021; Hsiao, 2021). For newcomers, this identity negotiation involves constant balancing: acquiring host-country language and cultural norms while preserving heritage and traditions. This ongoing process shapes psychological well-being, social relationships, and everyday experiences (Joubert & Sibanda, 2022; Kiramba & Oloo, 2023).

However, a broad body of scholarship provides robust evidence that LLM-assisted writing introduces a new complication. When Sourati and colleagues (2025) analyzed LLM text revision, they found that semantic content remains preserved while stylistic elements such as personal voice, cultural markers, and individual language choices undergo systematic alteration toward dominant patterns in training data. When a Senegalese immigrant writes in French with characteristic phrases, cultural references, and linguistic patterns reflecting her heritage, LLM "improvement" strips those markers away, homogenizing her writing toward standardized global norms (Sourati et al., 2025; Kuteeva & Andersson, 2024). Her voice does not merely change, it disappears.

This erasure becomes particularly consequential in academic contexts. While ChatGPT

enhances lexical complexity in non-native English speakers' writing, this linguistic "equalization" paradoxically erases distinctive voice and cultural expression (Lin et al., 2025). The writing appears professionally polished yet experientially inauthentic, and it becomes improved by external metrics while internally fractured. For newcomers already navigating tensions between heritage and integration, LLM-assisted "improvement" intensifies authenticity concerns because the communication is fluent but no longer recognizably their own (Gao, 2021; Kiramba & Oloo, 2023; Ozbek-Damar, 2025). The tools designed to facilitate integration simultaneously eliminate the linguistic markers through which cultural identity is expressed and transmitted (Sourati et al., 2025).

4.2 Linguistic Homogenization and the Erosion of Heritage Language Vitality

Recent scholarship indicates increasing concern regarding linguistic homogenization, which means the gradual erosion of diverse language and cultural expressions due to pressures to conform to dominant language norms (Rahmani & Karimi, 2025). In globalized and digital contexts, dominant languages like English increasingly overshadow minority and heritage languages, resulting in significant declines in linguistic diversity and cultural distinctiveness (Skutnabb-Kangas & May, 2017).

Large language models amplify this process by reinforcing standardized linguistic patterns. Sourati et al. (2025) demonstrate that texts revised by LLMs show diminished stylistic and cultural variation, as features reflecting non-dominant languages and individual expression are systematically integrated into dominant forms. Similarly, Milička et al. (2025) found through multidimensional analysis that LLM-generated texts exhibit reduced stylistic variation compared to human writing, with AI maintaining more consistent and thus more homogenized output across registers. This change limits the expressive capacity of communication and diminishes the cultural frameworks that communities use to interpret experiences (Lin et al., 2025; Kuteeva & Andersson, 2024).

Zeng and Yang (2024) contend that English hegemony, reinforced through AI systems predominantly trained on English data, marginalizes minority languages and epistemologies. Without targeted initiatives for linguistic pluralism, LLMs exacerbate English dominance, undermining cognitive and cultural diversity (Zeng & Yang, 2024; Li et al., 2024). Plum et al. (2025) argue LLMs lack "cultural reasoning," defined as the ability to recognize and adjust for culture-specific knowledge, values, and norms, which sustains stereotypes and ignores minority perspectives (Plum et al., 2025; Seth, 2025).

Heritage language speakers thus face a fundamental conflict: technological inclusion versus cultural identity preservation (Fenech-Borg et al., 2025). While LLMs possess cultural knowledge, they remain insensitive to cultural differences in practice, often requiring manual correction for appropriate adaptation (Singh et al., 2024; Tenzer et al., 2025).

4.3 Assimilation Pressure and Authenticity Concerns in Technology-Mediated Communication

Across the studies reviewed, there is consistent and accumulating evidence that newcomers encounter both overt and subtle pressures to adopt host-country communication norms, often at the expense of heritage language practices, idiomatic expressions, and culturally grounded ways of speaking (Alshihry, 2024; Karpava, 2024; Tenzer et al., 2025).

Research repeatedly shows that such pressures can prompt individuals to question whether their linguistic choices reflect their authentic identities or merely conform to institutional expectations to particularly in contexts where mastery of the dominant language is framed as essential for educational and professional inclusion (Migliarini, 2024; Marrone, 2017). Emerging work further indicates that when AI enters this landscape, institutional norms become intertwined with LLM-mediated communication tools, shaping not only what newcomers articulate but also how they understand themselves as they navigate between home and host cultures (García & Wei, 2014; Feng et al., 2025).

Authenticity is central to immigrants' language experiences: many wrestle with whether their speech or writing sounds like "themselves" or simply satisfies others' expectations (Alshihry, 2024; Karpava, 2021; Eerdemutu et al., 2024). LLMs intensify this tension. Experimental work shows that LLM outputs exhibit high semantic alignment with prompts but relatively low stylistic alignment, prioritizing content over individual style (Durandard et al., 2025). Studies of LLM-driven editing and lexical shifts similarly find that AI revisions preserve core meaning while converging on more standardized, high-prestige forms of expression (Lin et al., 2025; Milička et al., 2025). The result is language that reads as polished and professional but feels alien to its author.

For newcomers adapting to a new culture, this dynamic is especially consequential. AI-assisted communication may appear fluent and acceptable yet lack authenticity, creating psychological strain and a disconnect between communication and self (Bélanger & Verkuyten, 2023; Karpava, 2024). At precisely the time when immigrants need to maintain a connection to their heritage identity while developing competence in the dominant language, LLM-mediated revision can erode that link by stripping linguistic markers of identity and reinforcing assimilation pressures. In doing so, it risks undermining the very adaptation and inclusion processes it purports to support (Alshihry, 2024; Migliarini, 2024).

Section V: LLMs in Context—Distinguishing Tools and Understanding Mechanisms

5.1 Large Language Models versus Neural Machine Translation: A Critical Distinction

To understand the challenges LLMs pose in cross-cultural communication, they must be distinguished from neural machine translation (NMT) systems. NMT systems translate text between languages, seeking to preserve meaning accurately. Using encoder–decoder architectures, they learn relationships between language pairs and perform direct translations such as "buenos días" → "good morning" (Ye, 2025; Boukhari & Regedor, 2025). LLMs, by contrast, generate fluent, contextually appropriate responses. They learn statistical patterns across languages that incorporate cultural knowledge and communicative norms (Hu et al., 2024; Li et al., 2024; Liu et al., 2024; Sun et al. 2025). When answering cultural questions, they synthesize information from training data, which may embed bias. A newcomer using Google Translate to read a sign relies on a tool with predictable limits. Asking ChatGPT for cultural advice, however, engages a system that seems informed yet often reflects Western perspectives and lacks the nuanced understanding users expect.

5.2 Authenticity and Naturalness Concerns in AI-Mediated Communication

Authenticity in AI-mediated cross-cultural communication involves not only identity preservation but also questions of genuine understanding and interaction. These concerns arise at several levels, including information about culture, AI-assisted communication, and relationships mediated through technology.

A growing body of research shows that LLMs often rely on surface-level stereotypes rather than genuine understanding of cultural values (Kharchenko et al., 2025; Lawton & Ibarrola, 2023). They can produce fluent descriptions of cultural practices yet frequently reflect Western interpretations rather than local perspectives. For instance, when describing Japanese notions of honor or Mexican family structures, LLMs depend mainly on English-language sources written by Western observers instead of knowledge from within those cultures. As a result, they appear culturally knowledgeable but reproduce secondhand knowledge about cultures rather than from them. For newcomers trying to understand a host culture, this distinction is crucial because such information may be accurate yet still miss the nuance and lived complexity of real practice.

A related issue is naturalness in communication. LLM-assisted writing may appear fluent and standardized yet feel less authentic to the writer's own voice and cultural background (Hwang et al., 2025). This tension between fluency and authenticity continues to challenge newcomers seeking ways to communicate that are both effective and true to self.

Section VI: Institutional and Policy Responses—From Prohibition to Participation

6.1 Educational Institution Responses: Between Prohibition and Integration

It's well documented that educational institutions have adopted diverse strategies for using LLMs in cross-cultural contexts, shaped by institutional goals, cultural values, and teaching philosophies. Approaches range from outright bans to structured integration frameworks that balance AI's benefits with the need for autonomy and academic integrity (Barnes et al., 2024; Cotton et al., 2024; Gulumbe et al., 2025; Nnorom, 2025).

The prohibition approach addresses valid concerns about integrity, dependency, and skill loss. However, it is difficult to enforce and can disadvantage non-native speakers who rely on AI to achieve native-level writing (Yusuf et al., 2024). Such bans can harm students who need assistance without deterring those who use AI covertly.

More progressive institutions adopt disclosure frameworks requiring students to acknowledge AI use while ensuring equitable access across cultures and socioeconomic groups (Yusuf et al., 2024). These frameworks accept AI's ubiquity and shift the question from *whether* students will use it to *how*: either to support learning or replace cognitive engagement.

Nordic universities base AI policies on values of trust, transparency, and openness (Butt, 2024; Cannavale et al., 2025; Masso et al., 2024; Rekman, 2024). Cultural contexts shape these designs, emphasizing collaboration over punishment and autonomy over control. Elsewhere, institutions prioritize innovation, competitiveness, or cost efficiency, producing divergent policies despite shared technological challenges (Cai & Yin, 2025; Goffi & Momcilovic, 2022; Han et al., 2025; Hongladarom & Bandasak, 2024; Kochupillai et al., 2022; Kum et al., 2024; Núñez, 2025; Popa Tache & Vâlcu, 2025; Wong, 2025).

Research highlights the importance of culturally responsive policies and sustained faculty development (Al-Zahrani & Alasmari, 2024; Ahmed, 2024). Effective AI integration requires solid infrastructure, faculty training, institutional support, and commitment to educational quality and cultural equity. Institutions that invest in comprehensive faculty programs more effectively distinguish between AI uses that enhance learning and those that undermine integrity or reinforce bias against non-native and marginalized students (Ma et al., 2024).

6.2 Community Organization Strategies: Building Support Beyond Institutions

Community organizations address challenges in AI-mediated communication by developing guidelines for responsible, culturally aware AI use, building peer support networks to reduce dependence, and offering education on both benefits and risks (Salas-Pilco et al., 2022). These initiatives promote digital inclusivity and cultural literacy, helping immigrants, international students, and minorities navigate an increasingly AI-mediated world.

Research on community interventions highlights the need for multidimensional solutions that combine educational, technological, and social approaches (Salas-Pilco et al., 2022; Marko et al., 2025). Technology alone is insufficient. Effective interventions focus on three dimensions: pedagogical (media literacy and critical AI evaluation), technological (accessible, multilingual, culturally responsive tools), and sociocultural (addressing power dynamics, cultural hierarchies, and identity issues in AI use).

Peer support programs are key to reducing dependency. They address the psychological and social dimensions institutional policies often overlook by promoting group reflection, setting boundaries between AI assistance and independent learning, and providing emotional support for those concerned about identity in AI-mediated communication (Salas-Pilco et al., 2022). Community organizations also act as intermediaries between policymakers and local populations, advocating for solutions tailored to each community's needs rather than one-size-fits-all governance models.

6.3 Professional and Organizational Adaptations: Navigating Workplace Complexities

Employers are developing AI workplace policies that recognize the technology's ubiquity while upholding expectations for cross-cultural competence, authentic communication, and ethical conduct (Tang et al., 2023; Rakova et al., 2020). Research shows that organizational culture, leadership support, and continuous training are essential for AI integration that strengthens work quality and inclusion (Ahmad et al., 2023; Einola & Khoreva, 2022).

A major challenge lies in distinguishing productive AI use from harmful practices. The workslop phenomenon, which indicate low-quality AI-generated content, remains difficult to regulate (Rakova et al., 2020; Bankins et al., 2024). Sharing such content, especially across cultural and linguistic boundaries, increases colleagues' cognitive load, erodes trust, and risks perpetuating stereotypes. Many organizations still lack frameworks to differentiate between beneficial AI tools, such as translation and accessibility support, and workslop that adds little value (Bankins et al., 2024).

Effective responses require AI literacy that merges technical and cultural competence (Sienkiewicz-Małyjurek & Zyzak, 2024). Training should go beyond technical skills to include critical reflection on culturally sensitive AI use, the preservation of nuance, and transparent communication across cultures. Organizations that foster supportive cultures, model ethical AI use, and sustain employee development achieve greater success in using AI to enhance rather than diminish cross-cultural communication (Sienkiewicz-Małyjurek & Zyzak, 2024).

6.4 Cross-Cultural and Policy Considerations: Values, Disparities, and Governance

Global disparities in AI adoption and digital infrastructure undermine equitable cross-cultural communication. Studies reveal stark divides between high- and low-resource regions, where limited access, poor service quality, and weak institutional capacity hinder context-specific AI governance (Al-Zahrani & Alasmari, 2024; Ahmed, 2024; Marko et al., 2025). As a result, marginalized groups, including speakers of low-resource languages, residents of developing regions, immigrants, and refugees, struggle to access supportive AI tools while being overexposed to Western-trained systems that reinforce cultural bias (Ahmed, 2024). The compounding effect is a form of technological marginalization that mirrors and intensifies existing global inequalities.

Addressing these challenges requires context-specific policies attuned to local realities rather than universal models. Effective strategies must span four dimensions: technical (culturally responsive, accessible tools), governance (locally grounded frameworks), educational (digital literacy and critical AI skills), and equity (reducing global disparities in access and quality) (Abbasnejad et al., 2025; Abbasi et al., 2025; Kudriashova & Martynenko, 2025).

Discussion: Implications, Limitations, and Paths Forward

7.1 Synthesis of Major Findings

Taken together, the results show that cultural bias in large language models is not incidental but structural. At the architectural level, models are trained on corpora that overrepresent English and Western epistemologies, encoding Western-centric value priorities into their parameters and alignment processes even when outputs appear neutral. At the experiential level, newcomers encounter this bias as cultural misrepresentation, stereotype reinforcement, and communication-style mismatch: LLMs often recast culturally grounded reasoning into Western academic forms, mishandle idioms and metaphors, and recommend interaction norms misaligned with high-context or relational cultures.

These patterns translate into what the review terms an equity paradox. Non-native speakers face a cruel double bind in AI detection systems, where both authentic writing and AI-assisted improvement can trigger suspicion, while speakers of low-resource languages confront higher costs and lower-quality service. Workplace uses of LLMs can generate workslop that erodes trust and disproportionately harms those who depend on AI for linguistic support. In education, performance gaps, brittleness across languages, and misaligned benchmarks further disadvantage multilingual learners.

The findings also document processes of identity erasure. LLM-based rewriting preserves semantic content but systematically strips stylistic and cultural markers, pushing users toward standardized global norms and weakening heritage language vitality. For newcomers, this produces a tension between fluency and authenticity, as AI-assisted communication may read as acceptable yet feel detached from their sense of self. Distinguishing LLMs from more bounded tools like neural machine translation clarifies why these effects arise: LLMs do not simply translate but synthesize and normalize cultural knowledge.

Finally, the review shows that institutional and policy responses remain uneven. Educational, community, and workplace initiatives are beginning to address AI's role in cross-cultural communication, but many frameworks remain top-down and insufficiently participatory. Across studies, a consistent message emerges: mitigating harm and supporting equitable use will require governance, design, and community practices that explicitly center underrepresented users rather than treating LLMs as culturally neutral tools.

7.2 Theoretical Implications

LLMs extend identity-protective cognition to the technological sphere. Users interpret information through the lens of their group identity rather than rational evaluation; algorithms also embed group-centric values during training and deployment. LLMs amplify majority-culture patterns, creating asymmetries that advantage majority users while marginalizing others. Simultaneously, LLMs redefine authenticity, privileging communication styles aligned with dominant training data rather than with lived cultural experience.

7.3 Implications for Newcomers

Newcomers and immigrants face a compound burden: underrepresentation of their cultures in LLMs, inadequate service at high costs, and homogenization of expression that threatens cultural preservation. For vulnerable populations, LLMs pose risks with opportunities.

7.4 Policy and Governance Implications

Current governance frameworks lack participatory structures centered on affected communities. Effective governance requires treating AI as a question of power and representation, fundamentally political issues demanding democratic rather than purely technical solutions. Governance must enable data sovereignty and give underrepresented communities control over how their cultures are represented, ensuring collaboration rather than imposition.

7.5 Technical and Design Implications

Corpus curation must intentionally represent diverse cultural values and non-Western epistemologies. Community participation and data sovereignty are prerequisites, not optional. Evaluation must extend beyond technical accuracy to assess cultural appropriateness and the impact on equity, utilizing culturally diverse evaluators. Architectures must integrate transparency and community feedback rather than remain opaque black boxes.

7.6 Limitations and Future Research

This semi-systematic review strikes a balance between rigor and feasibility in a rapidly evolving field. Limitations include reliance on English-language sources, the 2020–2025 timeframe, limited longitudinal evidence on cultural identity formation, and underrepresentation of research from non-English-speaking regions and lower-income countries. Future research should examine identity development through longitudinal cross-cultural studies, evaluate community-based and technical interventions, and analyze how LLM bias intersects with other forms of marginalization.

7.7 Conclusion: Toward Technology that Serves Rather Than Dominates

LLM proliferation demands attention to equity, cultural preservation, and authentic understanding. These systems embed biases advantaging some groups while risking cultural erasure where diversity most needs protection. Our task is to determine which systems to build, whose interests they serve, and what values they express. AI governance is fundamentally about power and representation, requiring participatory structures that center on and represent affected communities.

LLMs cannot replace the human work of building cross-cultural understanding. For newcomers navigating integration, for communities protecting linguistic and cultural distinctiveness, and for those committed to equitable communication, the key question is not only what AI can do but what it should refrain from doing. The tools will remain; our task is to use them in ways that support, rather than displace, the slow and demanding work of genuine human connection.

References

Abbasi, B. N., Wu, Y., & Luo, Z. (2025). Exploring the impact of artificial intelligence on curriculum development in global higher education institutions. Education and Information Technologies, 30(1), 547-581.

Abbasnejad, B., Soltani, S., Taghizadeh, F., & Zare, A. (2025). Developing a multilevel framework for AI integration in technical and engineering higher education: insights from bibliometric analysis and ethnographic research. Interactive Technology and Smart Education.

Abdilla, A., & Crawford, K. (2020). Indigenous knowledge systems are not supplementary to Western paradigms. International Journal of Communication, 14, 1–12.

Agarwal, D., Naaman, M., & Vashistha, A. (2025, April). AI suggestions homogenize writing toward western styles and diminish cultural nuances. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (pp. 1-21).

Ahia, O., Kumar, S., Gonen, H., Kasai, J., Mortensen, D. R., Smith, N. A., & Tsvetkov, Y. (2023). Do all languages cost the same? tokenization in the era of commercial language models. arXiv preprint arXiv:2305.13707.

Ahmad, S. F., Han, H., Alam, M. M., Rehmat, M., Irshad, M., Arraño-Muñoz, M., & Ariza-Montes, A. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. Humanities and Social Sciences Communications, 10(1), 1-14.

Ahmed, S. (2024). Beyond Human Teachers: Post-Humanist Perspectives on AI, Cultural Inequities, and Educational Transformation. Journal of Posthumanism, 4(3), 364-373.

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. Humanities and

Social Sciences Communications, 11(1), 1-12.

Algouzi, S., & Alzubi, A. A. F. (2023). The study of AI-mediated communication and socio-cultural language-related variables: Gmail reply suggestions. Applied Artificial Intelligence, 37(1), 2175114.

Allan, K., Azcona, J., Sripada, S., Leontidis, G., Sutherland, C. A., Phillips, L. H., & Martin, D. (2025). Stereotypical bias amplification and reversal in an experimental model of human interaction with generative artificial intelligence. Royal Society Open Science, 12(4), 241472.

Alshihry, M. A. (2024). Heritage language maintenance among immigrant youth: Factors influencing proficiency and identity. In S. Karpava (Ed.), Heritage language policy in early childhood education (pp. 45–65). Springer.

Alyafeai, Z., Alshaibani, M. S., AlKhamissi, B., Luqman, H., Alareqi, E., & Fadel, A. (2023). Taqyim: Evaluating arabic nlp tasks using chatgpt models. arXiv preprint arXiv:2306.16322.

Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., … & Veríssimo, D. (2023). The manifold costs of being a non-native English speaker in science. PLoS biology, 21(7), e3002184.

Arıcı, F. (2024). Examination of research conducted on the use of artificial intelligence in science education. Sakarya University Journal of Education, 14(3), 539-562.

Asfahani, A. M. (2022). The impact of artificial intelligence on industrial-organizational psychology: A systematic review. The Journal of Behavioral Science, 17(3), 125-139.

Ashraf, Y., Wang, Y., Gu, B., Nakov, P., & Baldwin, T. (2025, April). Arabic dataset for LLM safeguard evaluation. In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 5529-5546).

Asseri, B., Abdelaziz, E., & Al-Wabil, A. (2025). Prompt Engineering Techniques for Mitigating Cultural Bias Against Arabs and Muslims in Large Language Models: A Systematic Review. arXiv preprint arXiv:2506.18199.

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. Journal of organizational behavior, 45(2), 159-182.

Bao, T., Zhao, Y., Mao, J., & Zhang, C. (2025). Examining linguistic shifts in academic writing before and after the launch of ChatGPT: a study on preprint papers. Scientometrics, 1-31.

Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. Science, Technology, & Human Values, 47(5), 855-881.

Barnes, A. J., Zhang, Y., & Valenzuela, A. (2024). AI and culture: Culturally dependent responses to AI systems. Current Opinion in Psychology, 58, 101838.

Bélanger, E., & Verkuyten, M. (2023). Language and belonging among older immigrants. Journal of Ethnic and Migration Studies, 49(16), 3973–3992. https://doi.org/10.1080/1369183X.2023.2280994

Bignotti, F. (2025). Potential bias in AI: cultural representation and the marginalization of African art.

Binkyte, R. (2025). Interactional Fairness in LLM Multi-Agent Systems: An Evaluation Framework. arXiv preprint arXiv:2505.12001.

Birhane, A., Isaac, W., Prabhakaran, V., Diaz, M., Elish, M. C., Gabriel, I., & Mohamed, S. (2022, October). Power to the people? Opportunities and challenges for participatory AI. In Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (pp. 1-8).

Björk, B. C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. Journal of informetrics, 7(4), 914-923.

Borah, A., & Mihalcea, R. (2024). Towards implicit bias detection and mitigation in multi-agent llm interactions. arXiv preprint arXiv:2410.02584.

Boukhari, O., & Regedor, A. M. (2025). Language as a Barrier, Bridge and Battleground: A Study of Linguistic Experiences in the Context of Moroccan Migration to Southern Portugal. International Journal of Linguistics, Literature and Translation, 8(5), 727-742.

Brandt, B. (2023). Replicating human bias through synthetic data generation using deep learning (Doctoral dissertation, University of Wisconsin-Whitewater).

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889-942.

Butt, J. S. (2024). A comparative study about the use of artificial intelligence (AI) in public administration of Nordic states with other European economic sectors. Euro Economica, 43(1), 40-66.

Cai, C., & Yin, J. (2025). Cultural and ethical foundations of AI governance divergence: a comparative analysis of China and the west. Revista Política Internacional, 7(1), 215-233.

Cannavale, C., Claudio, L., & Koroleva, D. (2025). Digitalisation and artificial intelligence development. A cross-country analysis. European Journal of Innovation Management, 28(11), 112-130.

Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). How people use chatgpt (No. w34255). National Bureau of Economic Research.

Chiu, Y. Y., Jiang, L., Lin, B. Y., Park, C. Y., Li, S. S., Ravi, S., Bhatia, M., Antoniak, M., Tsvetkov, Y., Shwartz, V., & Choi, Y. (2024). CulturalBench: A robust, diverse and challenging benchmark on measuring the (lack of) cultural knowledge of LLMs. arXiv preprint. https://arxiv.org/abs/2410.02677

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. Annals of internal medicine, 126(5), 376-380.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in education and teaching international, 61(2), 228-239.

Crystal, D. (2003). English as a global language. Cambridge university press.

Dairo, D. Navigating Cultural Diversity in the Digital Age: Legal Perspectives on Opportunities and Challenges of Artificial Intelligence. In Cultural Odyssey: 20 Years of Implementation of UNESCO's 2005 Convention in Nigeria (p. 181).

Darginavičienė, I. (2023). The multilingualism: Language and cultural identity. LOGOS-A Journal of Religion, Philosophy, Comparative Cultural Studies and Art, (116), 167-174.

De Sousa Santos, B. (2014). Epistemologies of the South: Justice against epistemicide. Paradigm Publishers.

Durandard, N., Dhawan, S., & Poibeau, T. (2025). LLMs stick to the point, humans to style: Semantic and stylistic alignment in human and LLM communication. Proceedings of the 26th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 206–213.

Durmus, E., Nguyen, K., Liao, T. I., Schiefer, N., Askell, A., Bakhtin, A., … & Ganguli, D. (2023). Towards measuring the representation of subjective global opinions in language models. arXiv preprint arXiv:2306.16388.

Eger, S., Cao, Y., D'Souza, J., Geiger, A., Greisinger, C., Gross, S., … & Miller, T. (2025). Transforming science with large language models: A survey on ai-assisted scientific discovery, experimentation, content generation, and evaluation. arXiv preprint arXiv:2502.05151.

Ehrensberger-Dow, M., Albl-Mikasa, M., Andermatt, K., Hunziker Heeb, A., & Lehr, C. (2020). Cognitive load in processing ELF: Translators, interpreters, and other multilinguals. Journal of English as a lingua franca, 9(2), 217-238.

Einola, K., & Khoreva, V. (2023). Best friend or broken tool? Exploring the co-existence of humans and artificial intelligence in the workplace ecosystem. Human Resource Management, 62(1), 117-135.

Elgamal, A. M. A. (2019). Cognitive Factors Affecting Language Learning and Acquisition of Native and Non-Native Speakers. Journal of Research in Curriculum Instruction and Educational Technology, 4(4), 135-152.

Fenech-Borg, E. Z., Meznaric-Kos, T. P., Lekovic-Bojovic, M. D., & Hentze-Djurhuus, A. J. (2025). The Cultural Gene of Large Language Models: A Study on the Impact of Cross-Corpus Training on Model Values and Biases. arXiv preprint arXiv:2508.12411.

Feng, H., Li, K., & Zhang, L. J. (2025). What does AI bring to second language writing? A systematic review (2014-2024). Language Learning & Technology, 29(1).

Ferdaus, M. M., Abdelguerfi, M., Loup, E., N. Niles, K., Pathak, K., & Sloan, S. (2024). Towards trustworthy ai: A review of ethical and robust large language models. ACM Computing Surveys.

Ferré, P., Fraga, I., & Hinojosa, J. A. (2025). The interplay between language and emotion: a narrative review. Cognition and Emotion, 1-28.

Fielding, R. (2021). A multilingual identity approach to intercultural stance in language learning. The Language Learning Journal, 49(4), 466-482.

Fock, A., & Siller, H. S. (2025). Generative Artificial Intelligence in Secondary STEM Education in the Light of Human Flourishing: A Scoping Literature Review.

Fraser, K. C., Dawkins, H., & Kiritchenko, S. (2025). Detecting ai-generated text: Factors influencing detectability with current methods. Journal of Artificial Intelligence Research, 82, 2233-2278.

Gao, F. (2021). Negotiation of native linguistic ideology and cultural identities in English learning: a cultural schema perspective. Journal of Multilingual and Multicultural Development, 42(6), 551-564.

Garcia, M. B. (2025). ChatGPT as an Academic Writing Tool: Factors Influencing Researchers' Intention to Write Manuscripts Using Generative Artificial Intelligence. International Journal of Human–Computer Interaction, 1-15.

García, O., & Wei, L. (2014). Translanguaging: Language, bilingualism and education. Palgrave Macmillan.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. Proceedings of the National Academy of Sciences, 115(16), E3635–E3644. https://doi.org/10.1073/pnas.1720347115

Ghimire, A. (2025). Utilizing ChatGPT to integrate world English and diverse knowledge: A transnational perspective in critical artificial intelligence (AI) literacy. Computers and Composition, 75, 102913.

Glickman, M., & Sharot, T. (2025). How human–AI feedback loops alter human perceptual, emotional and social judgements. Nature Human Behaviour, 9(2), 345-359.

Goffi, E., & Momcilovic, A. (2022). Respecting cultural diversity in ethics applied to AI: A new approach for a multicultural governance. Misión Jurídica, 15(23), 111-122.

Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. Journal of family medicine and primary care, 2(1), 9-14.

Gulumbe, B. H., Audu, S. M., & Hashim, A. M. (2025). Balancing AI and academic integrity: What are the positions of academic publishers and universities?. AI & SOCIETY, 40(3), 1775-1784.

Guo, J., & Xu, Y. (2025). Your AI Bosses Are Still Prejudiced: The Emergence of Stereotypes in LLM-Based Multi-Agent Systems. arXiv preprint arXiv:2508.19919.

Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., & de Beer, J. (2020). Artificial intelligence (AI) deployments in Africa: Benefits, challenges and policy dimensions. The African Journal of Information and Communication, 26, 1–28. https://doi.org/10.23962/10539/30361

Haase, J., & Pokutta, S. (2025). Beyond Static Responses: Multi-Agent LLM Systems as a New Paradigm for Social Science Research. arXiv preprint arXiv:2506.01839.

Han, X., HL Kaas, M., & Wang, C. D. (2025). A cross-cultural examination of fairness beliefs in human-AI interaction. Han, X., Kass, M., & Wang, C.(forthcoming). A cross-cultural examination of fairness beliefs in human-AI interaction. In Ethics of Institutional Beliefs: From Theoretical to Empirical. Edward Elgar Publishing.

Hamann, S. A., Giang, J., De Maeseneer, M. G., Nijsten, T. E., & van den Bos, R. R. (2017). Editor's choice–five year results of great saphenous vein treatment: a meta-analysis. *European journal of vascular and endovascular surgery*, *54*(6), 760-770.

Havaldar, S., Cho, Y. M., Rai, S., & Ungar, L. (2025, November). Culturally-Aware Conversations: A Framework & Benchmark for LLMs. In Proceedings of the Fourth Workshop on Bridging Human-Computer Interaction and Natural Language Processing (HCI+ NLP) (pp. 220-229).

Hergert, L., Berend, G., Szegedy, M., Turan, G., & Jelasity, M. (2025). On the Brittleness of LLMs: A Journey around Set Membership. arXiv preprint arXiv:2511.12728.

Hofmann, V., Kalluri, P. R., Jurafsky, D., & King, S. (2024). AI generates covertly racist decisions about people based on their dialect. Nature. https://doi.org/10.1038/s41586-024-07856-5

Hongladarom, S., & Bandasak, J. (2024). Non-western AI ethics guidelines: Implications for intercultural ethics of technology. Ai & Society, 39(4), 2019-2032.

Hoppers, O. C. (2002). Indigenous knowledge and the integration of knowledge systems. Indigenous knowledge and the integration of knowledge systems: Towards a philosophy of articulation, 2-22.

Hsiao, C. Y. (2021). Online and Offline Adaptation among Transnational Newcomers: Technology-mediated Social Exchange and Trust Development (Doctoral dissertation).

Hu, H., Zhou, Q., & Hashim, H. (2025). Negotiating identity in the age of ChatGPT: non-native English researchers' experiences with AI-assisted academic writing. Humanities and Social Sciences Communications, 12(1), 1-11.

Hu, X., Zhang, F., Chen, S., & Yang, Z. (2024). Unveiling the statistical foundations of chain-of-thought prompting methods. arXiv preprint arXiv:2408.14511.

Hwang, A. H. C., Liao, Q. V., Blodgett, S. L., Olteanu, A., & Trischler, A. (2025). 'It was 80% me, 20% AI': Seeking Authenticity in Co-Writing with Large Language Models. Proceedings of the ACM on Human-Computer Interaction, 9(2), 1-41.

Jiang, Y., Hao, J., Fauss, M., & Li, C. (2024). Detecting ChatGPT-generated essays in a large-scale writing assessment: Is there a bias against non-native English speakers?. Computers & Education, 217, 105070.

Jiang, Y., Zhao, J., Yuan, Y., Zhang, T., Huang, Y., Zhang, Y., … & Li, X. (2025). Never compromise to vulnerabilities: A comprehensive survey on ai governance. arXiv preprint arXiv:2508.08789.

Joubert, M., & Sibanda, B. (2022). Whose language is it anyway? Students' sense of belonging and role of English for higher education in the multilingual, South African context. South African Journal of Higher Education, 36(6), 47-66.

Kamran, F. (2024). Decolonizing artificial intelligence: Indigenous knowledge and digital epistemic justice. Journal of Critical AI Studies, 2(1), 45–68.

Karpava, S. (2021). Heritage language use, maintenance and transmission by second-generation immigrants in Cyprus. In P. Romanowski & M. Jedynak (Eds.), Current Research in Bilingualism and Education (pp. 83–108). Springer.

Karpava, S. (2024). The hybrid linguistic and cultural identity of second-generation immigrants in Cyprus. In S. Karpava (Ed.), Hybrid linguistic and cultural identities in migration contexts (pp. 101–124). Routledge.

Keleg, A. (2025). LLM Alignment for the Arabs: A Homogenous Culture or Diverse Ones?. arXiv preprint arXiv:2503.15003.

Keller, J. D., & Potthast, R. (2024). AI-based data assimilation: Learning the functional of analysis estimation. arXiv preprint arXiv:2406.00390.

Khan, R., Qamar, M. T., Ansari, M. S., & Yasmeen, J. (2025). Enhancing or impairing? Exploring Indian EFL learners' academic writing narratives with ChatGPT. Cogent Education, 12(1), 2514329.

Kharchenko, J., Roosta, T., Chadha, A., & Shah, C. (2024). How well do llms represent values across cultures? empirical analysis of llm responses based on hofstede cultural dimensions. arXiv preprint arXiv:2406.14805.

Kiramba, L. K., & Oloo, J. A. (2023). "It's OK. She doesn't even speak English": Narratives of language, culture, and identity negotiation by immigrant high school students. Urban Education, 58(3), 398-426.

Klassen, T. P., Lawson, M. L., & Moher, D. (2005). Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. Journal of clinical epidemiology, 58(8), 769-776.

Kochupillai, M., Kahl, M., Schmitt, M., Taubenböck, H., & Zhu, X. X. (2022). Earth observation and artificial intelligence: Understanding emerging ethical issues and opportunities. IEEE Geoscience and Remote Sensing Magazine, 10(4), 90-124.

Koo, W. W. (2025). Cross-lingual effects of AI-generated content on human work. Scientific Reports, 15(1), 30949.

Kum, H. C., Bedrick, S., & Weigle, M. C. (2024). Challenges in Data Science. Digital Ethology: Human Behavior in Geospatial Context, 33, 211.

Kuteeva, M., & Andersson, M. (2024). Diversity and Standards in Writing for Publication in the Age of AI—Between a Rock and a Hard Place. Applied Linguistics, 45(3), 561-567.

Kwak, Y., & Pardos, Z. A. (2024). Bridging large language model disparities: Skill tagging of multilingual educational content. British Journal of Educational Technology, 55(5), 2039-2057.

Lawton, T., Grace, K., & Ibarrola, F. J. (2023, July). When is a tool a tool? user perceptions of system agency in human–ai co-creative drawing. In Proceedings of the 2023 ACM Designing Interactive Systems Conference (pp. 1978-1996).

Lee, E., & Moshirnia, A. (2024). The AI Penalty: Is There a Bias against AI-Generated Works?. Mich. St. L. Rev., 641.

Lege, R. P. Auditing the Fairness of AI-Detection Tools: A Comparative Study of ESL, Published, and AI-Generated Texts and Their Misclassification Risks. International Journal of Teaching, Learning and Education, 4(5), 638273.

Lehdonvirta, V. (2022). Cloud empires: How digital platforms are overtaking the state and how we can regain control. Mit press.

Li, C., Chen, M., Wang, J., Sitaram, S., & Xie, X. (2024). Culturellm: Incorporating cultural differences into large language models. Advances in Neural Information Processing Systems, 37, 84799-84838.

Li, J., Papay, S., & Klinger, R. (2025). Are Humans as Brittle as Large Language Models?. arXiv preprint arXiv:2509.07869.

Li, Y., Huang, Y., Wang, H., Cheng, Y., Zhang, X., Zou, J., & Sun, L. Evaluating Large Language Models with Psychometrics. In Large Language Models for Scientific and Societal Advances.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7).

Liang, W., Zhang, Y., Wu, Z., Lepp, H., Ji, W., Zhao, X., … & Zou, J. (2025). Quantifying large language model usage in scientific papers. Nature Human Behaviour, 1-11.

Lin, D., Zhao, N., Tian, D., & Li, J. (2025). ChatGPT as Linguistic Equalizer? Quantifying LLM-Driven Lexical Shifts in Academic Writing. arXiv preprint arXiv:2504.12317.

Lin, Z., & Zhao, X. Cultural Tailoring Paradox: Navigating Perceived Homophily and AI Bias in Generative AI-Mediated Communication Among Black Communities. Available at SSRN 5261067.

Liu, X., Wu, Z., Wu, X., Lu, P., Chang, K. W., & Feng, Y. (2024). Are llms capable of data-based statistical and causal reasoning? benchmarking advanced quantitative reasoning with data. arXiv preprint arXiv:2402.17644.

Liu, Z., Zhang, J., Jiang, H., You, W., Pan, Y., Xu, S., … & Liu, T. (2025). AImanities and Mirror of Collectivized Mind: Philosophy Theories of Large Language Models.

Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. Smart Learning Environments, 11(1), 50.

Lodoen, S., & Orchard, A. (2025). Ethics and Persuasion in Reinforcement Learning from Human Feedback: A Procedural Rhetorical Approach. arXiv preprint arXiv:2505.09576.

Lyu, Y., & Du, Y. (2025). The ethical evaluation of large language models and its optimization. AI and Ethics, 1-14.

Ma, D., Akram, H., & Chen, I. H. (2024). Artificial intelligence in higher education: A cross-cultural examination of students' behavioral intentions and attitudes. International Review of Research in Open and Distributed Learning, 25(3), 134-157.

Mao, R., Liu, Q., Li, X., Cambria, E., & Hussain, A. (2025). Bridging Minds and Machines: Toward an Integration of AI and Cognitive Science. arXiv preprint arXiv:2508.20674.

Marko, J. G. O., Neagu, C. D., & Anand, P. B. (2025). Examining inclusivity: the use of AI and diverse populations in health and social care: a systematic review. BMC Medical Informatics and Decision Making, 25(1), 57.

Marrone, G. (2017). Linguistic and cultural assimilation as a human capital process. IZA Journal of Migration, 6(1), 1–27.

Masso, A., Kaun, A., & Van Noordt, C. (2024). Basic values in artificial intelligence: comparative factor analysis in Estonia, Germany, and Sweden. AI & society, 39(6), 2775-2790.

Mehdizadeh, A., & Hilbert, M. (2025). Homophily-induced emergence of biased structures in LLM-based multi-agent AI systems. Social Network Analysis and Mining, 15(1), 1-25.

Migliarini, V. (2024). Performing the good (im)migrant: Inclusion and expectations of (dis)abled asylum-seeking students. International Journal of Inclusive Education, 28(12), 1432–1451.

Milička, J., Marklová, A., & Cvrček, V. (2025). Benchmark of stylistic variation in LLM-generated texts. Corpus Linguistics and Linguistic Theory, 21(2), 255–283.

Naous, T., Laban, P., Xu, W., & Neville, J. (2025). Flipping the Dialogue: Training and Evaluating User Language Models. *arXiv preprint arXiv:2510.06552*.

Nguyen, J. K. (2024). Human bias in AI models? Anchoring effects and mitigation strategies in large language models. Journal of Behavioral and Experimental Finance, 43, 100971. https://doi.org/10.1016/j.jbef.2024.100971

Nguyen, X.-P., Aljunied, M., Joty, S., & Bing, L. (2024). Democratizing LLMs for low-resource languages by leveraging their English dominant abilities with linguistically-diverse prompts. Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 3501–3516. https://doi.org/10.18653/v1/2024.acl-long.192

Niederhoffer, K., Kellerman, G. R., Lee, A., Liebscher, A., Rapuano, K., & Hancock, J. T. (2025). AI-Generated "Workslop" Is Destroying Productivity. Harvard Business Review.

Nnorom, I. C. (2025). Ethical Considerations in Artificial Intelligence and Academic Integrity: Balancing Technology and Human Values. AI and Ethics, Academic Integrity and the Future of Quality Assurance in Higher Education, 15.

Núñez, S. H. (2025). Technology Transfer to Latin American Countries: Drifting Away from the United States and China?. Taylor & Francis.

Ozbek-Damar, S. (2025). Second Language (L2)–Speaker Immigrant Women's Perspectives on Identity Construction and L2 Socialization in a Community in the Southwest US(Doctoral dissertation, Arizona State University).

Öztürk, D. (2021). What does artificial intelligence mean for organizations? A systematic review of organization studies research and a way forward. The Impact of Artificial Intelligence on Governance, Economics and Finance, Volume I, 265-289.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. bmj, 372.

Pakray, P., Gelbukh, A., & Bandyopadhyay, S. (2025). Natural language processing applications for low-resource languages. Natural Language Processing, 31(2), 183-197.

Pareek, A. (2025). A Multi-Component AI Framework for Computational Psychology: From Robust Predictive Modeling to Deployed Generative Dialogue. arXiv preprint arXiv:2510.21720.

Pedersen, B. S., Sørensen, N., Nimb, S., Hansen, D. H., Olsen, S., & Al-Laith, A. (2025, March). Evaluating llm-generated explanations of metaphors–a culture-sensitive study of danish. In Proceedings of the Joint 25th Nordic Conference on Computational Linguistics and 11th Baltic Conference on Human Language Technologies (NoDaLiDa/Baltic-HLT 2025) (pp. 470-479).

Pérez Torres, M., Couso Lagarón, D., & Marquez Bargalló, C. (2023). Evaluation of STEAM project-based learning (STEAM PBL) instructional designs from the STEM practices perspective. *Education Sciences*, *14*(1), 53.

Peters, U., & Carman, M. (2024). Cultural bias in explainable AI research: A systematic analysis. Journal of Artificial Intelligence Research, 79, 971-1000.

Petrov, A., La Malfa, E., Torr, P., & Bibi, A. (2023). Language model tokenizers introduce unfairness between languages. Advances in neural information processing systems, 36, 36963-36990.

Plum, A., Lutgen, A. M., Purschke, C., & Rettinger, A. (2025). Identity-Aware Large Language Models require Cultural Reasoning. arXiv preprint arXiv:2510.18510.

Popa Tache, C. E., & Vâlcu, E. N. (2025). Artificial Intelligence and Corporate Liability Towards a New Legal-Ethical Contract in the Dynamics of Emerging Global Human Rights Convergences. Jurid. Trib.-Rev. Compar. &Int'l L., 15, 281.

Prakash, A., Aggarwal, S., Varghese, J. J., & Varghese, J. J. (2025). Writing without borders: AI and cross-cultural convergence in academic writing quality. Humanities and Social Sciences

Communications, 12(1), 1-11.

Qadri, R., Diaz, M., Wang, D., & Madaio, M. (2025). The case for" thick evaluations" of cultural representation in ai. arXiv preprint arXiv:2503.19075.

Rahman, A., Bowlin, G., Mohanty, B., & McGunigal, S. (2024). Towards Linguistically-Aware and Language-Independent Tokenization for Large Language Models (LLMs). arXiv preprint arXiv:2410.03568.

Rakova, B., Yang, J., Cramer, H., & Chowdhury, R. (2021). Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1-23.

Rao, P. S. (2019). The role of English as a global language. Research journal of English, 4(1), 65-79.

Rekman, O. (2024). Nordic Ethical AI Expert Group: Policy Recommendations for Ethical and Responsible AI.

Reusens, M., Kopetzky, D., & Buitelaar, P. (2024). Anchoring effects in LLM response quality: Systematic performance degradation for non-native English speakers. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 1–15).

Richardson, N., & Antonello, M. (2022). People at work 2022: A global workforce view. ADP Research Institute, 17, 2023.

Robertson, C., & Wolff, P. (2025). Llm world models are mental: Output layer evidence of brittle world model use in llm mechanical reasoning. arXiv preprint arXiv:2507.15521.

Rodrigues, L., Xavier, C., Costa, N., Batista, H., Silva, L. F. B., Chaleghi de Melo, W., … & Ferreira Mello, R. (2025, March). LLMs Performance in Answering Educational Questions in Brazilian Portuguese: A Preliminary Analysis on LLMs Potential to Support Diverse Educational Needs. In Proceedings of the 15th International Learning Analytics and Knowledge Conference (pp. 865-871).

Sætra, H. S. (2024). Preprinting in AI Ethics: Towards a Set of Community Guidelines. Available at SSRN 4598223.

Sahebi, S., & Formosa, P. (2025). The AI-mediated communication dilemma: epistemic trust, social media, and the challenge of generative artificial intelligence. Synthese, 205(3), 1-24.

Salas-Pilco, S. Z., Xiao, K., & Oshima, J. (2022). Artificial intelligence and new technologies in inclusive education for minority students: A systematic review. Sustainability, 14(20), 13572.

Sallam, M., Al-Adwan, A. S., Mijwil, M. M., Abdelaziz, D. H., Al-Qaisi, A., Ibrahim, O. M., & Sallam, M. (2025). Technology Readiness, Social Influence, and Anxiety as Predictors of University Educators' Perceptions of Generative AI Usefulness and Effectiveness.

Sallam, M., & Mousa, D. (2024). Evaluating ChatGPT performance in Arabic dialects: A comparative study showing defects in responding to Jordanian and Tunisian general health prompts. Mesopotamian Journal of Artificial Intelligence in Healthcare, 2024, 1-7.

Sanguinetti, P., & Palomo, B. (2025). Bridging Gaps in AI Representation: A Cross-Cultural Analysis of Media Imagery. Journalism Practice, 1-21.

Sato, K., Kaneko, H., & Fujimura, M. (2024). Reducing cultural hallucination in non-english languages via prompt engineering for large language models. OSF Preprints, 10.

Savadori, L., Dickson, M. M., Micciolo, R., & Espa, G. (2022). The polarizing impact of numeracy, economic literacy, and science literacy on the perception of immigration. Plos one, 17(10), e0274680.

Segerer, R. (2025). Cultural Value Alignment in Large Language Models: A Prompt-based Analysis of Schwartz Values in Gemini, ChatGPT, and DeepSeek. arXiv preprint arXiv:2505.17112.

Seth, A. (2025). Cultural Variability and Bias in Online Social Interactions and Large Language Models (Doctoral dissertation).

Shan, S., Li, Y., & Zhou, W. (2024). Cross-cultural implications of LLM deployment: Extended comparative analysis. International Journal of Human-Computer Studies, 185, 103198.

Shen, H., Clark, N., & Mitra, T. (2025, November). Mind the Value-Action Gap: Do LLMs Act in Alignment with Their Values?. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (pp. 3097-3118).

Singh, P., Patidar, M., & Vig, L. (2024). Translating across cultures: Llms for intralingual cultural adaptation. arXiv preprint arXiv:2406.14504.

Solatorio, A. V., Macalaba, R., & Liounis, J. (2025). Large Language Models and Synthetic Data for Monitoring Dataset Mentions in Research Papers. arXiv preprint arXiv:2502.10263.

Sourati, Z., Karimi-Malekabadi, F., Ozcan, M., McDaniel, C., Ziabari, A., Trager, J., … & Dehghani, M. (2025). The shrinking landscape of linguistic diversity in the age of large language models. arXiv preprint arXiv:2502.11266.

Su, J., Zhang, J., Ullrich, K., Bottou, L., & Ibrahim, M. (2025). A single character can make or break your LLM evals. arXiv preprint arXiv:2510.05152.

Suh, S., Bang, J., & Han, J. W. (2025). The Shift Towards Preprints in AI Policy Research: A Comparative Study of Preprint Trends in the US, Europe, and South Korea. arXiv preprint arXiv:2505.03835.

Sumartana, I. M., Hudiananingsih, P. D., & Rouf, M. A. (2025). Balancing globalization and linguistic heritage involves preserving mother tongues through inclusive education that values cultural identity and language diversity. Journal of Language, Literature, Social and Cultural Studies, 3(2), 179-196.

Sun, M., Han, R., Jiang, B., Qi, H., Sun, D., Yuan, Y., & Huang, J. (2025). A survey on large language model-based agents for statistics and data science. The American Statistician, 1-14.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research*, *104*, 333-339.

Tafa, T. O., Hashim, S. Z. M., Othman, M. S., Alhussian, H., Nasser, M., Abdulkadir, S. J., … & Bena, Y. A. (2025). Machine Translation Performance for LowResource Languages: A Systematic Literature Review. IEEE Access.

Tang, P., Koopman, J., Mai, K., De Cremer, D., Zhang, J. H., Reynders, P., Ng, C. T. S., & Chen, I. (2023). No person is an island: Unpacking the work and after-work consequences of interacting with artificial intelligence. Journal of Applied Psychology, 108(9), 1245–1270.

Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2024). Cultural bias and cultural alignment of large language models. PNAS Nexus, 3(9), pgae346. https://doi.org/10.1093/pnasnexus/pgae346

Tennant, J., Bauin, S., James, S., & Kant, J. (2018). The evolving preprint landscape: Introductory report for the Knowledge Exchange working group on preprints.

Tenzer, H., Abidi, O., & Feuerriegel, S. (2025). Designing LLMs for cultural sensitivity: Evidence from English-Japanese translation. arXiv preprint arXiv:2509.11921.

Tong, J., Sun, Y., Hubbard, R. A., Saine, M. E., Xu, H., Zuo, X., … & Chen, Y. (2025). Incorporating preprints in systematic reviews: a preliminary study of a novel method for rapid evidence synthesis. Journal of the American Medical Informatics Association, 32(11), 1654-1663.

UNESCO, T. (2021). UNESCO recommendation on open science. United Nations Educational, Scientific and Cultural Organization.

Urbaite, G. (2025). AI-Mediated English: How Generative Systems Reinforce English as a Global Lingua Franca. Porta Universorum, 1(9), 35-50.

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. Trends in cognitive sciences, 22(3), 213-224.

Van Bavel, J. J., Rathje, S., Vlasceanu, M., & Pretus, C. (2024). Updating the identity-based model of belief: From false belief to the spread of misinformation. Current Opinion in Psychology, 56, 101787.

Vasista, I., Mirza, I., Huang, C., Patil, R. R., Akalin, A., Zhu, K., & O'Brien, S. MALIBU Benchmark: Multi-Agent LLM Implicit Bias Uncovered. In ICLR 2025 Workshop on Building Trust in Language Models and Applications.

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. Technology, Knowledge and Learning, 1-22.

Wang, J. Y., Sukiennik, N., Li, T., Su, W., Hao, Q., Xu, J., … & Li, Y. (2024). A survey on human-centric llms. arXiv preprint arXiv:2411.14491.

Wang, K., Lyu, T., Su, G., Geiping, J., Yin, L., Canini, M., & Liu, S. (2025). When Fewer Layers Break More Chains: Layer Pruning Harms Test-Time Scaling in LLMs. arXiv preprint arXiv:2510.22228.

Wang, S., Xu, T., Li, H., Zhang, C., Liang, J., Tang, J., … & Wen, Q. (2024). Large language models for education: A survey and outlook. arXiv preprint arXiv:2403.18105.

Wong, P. H. (2025). Global Governance of AI, Cultural Values, and Human Rights. A Companion to Applied Philosophy of AI, 359-371.

Wu, A., Kuang, K., Zhu, M., Wang, Y., Zheng, Y., Han, K., … & Zhang, K. (2024). Causality for large language models. arXiv preprint arXiv:2410.15319.

Yang, R., Tong, J., Wang, H., Huang, H., Hu, Z., Li, P., … & Hong, C. (2025). Enabling inclusive systematic reviews: incorporating preprint articles with large language model-driven evaluations. Journal of the American Medical Informatics Association, 32(11), 1718-1725.

Yazar, B. K., Şahın, D. Ö., & Kiliç, E. (2023). Low-resource neural machine translation: A systematic literature review. IEEE Access, 11, 131775-131813.

Ye, X. (2025). Digital technologies and identity negotiation: a study of trilingual Uyghur university students' language learning experiences in intranational migrations. ReCALL, 37(2), 232-249.

Zangana, H. M., Amelia, P., Mustafa, F. M., & Li, S. (2025). The Role of English Language and AI in Scientific Writing: Ethical and Academic Implications. In Ensuring Secure and Ethical STM Research in the AI Era (pp. 191-218). IGI Global Scientific Publishing.

Zeng, J., & Yang, J. (2024). English language hegemony: retrospect and prospect. Humanities and Social Sciences Communications, 11(1), 1-9.

Zeng, W., Zhu, H., Qin, C., Wu, H., Cheng, Y., Zhang, S., … & Xiong, H. (2025). Application-Driven Value Alignment in Agentic AI Systems: Survey and Perspectives. arXiv preprint arXiv:2506.09656.

Zhang, W., Kam-Kwai, W., Xu, B., Ren, Y., Li, Y., Feng, Y., … & Chen, W. (2025, October). Cultiverse: Towards cross-cultural understanding for paintings with large language model. In Proceedings of the 33rd ACM International Conference on Multimedia (pp. 6710-6719).

Zheng, W. (2024). AI vs. Human: A Comparative Study of Cohesion and Coherence in Academic Texts between Human-Written and ChatGPT-Generated Texts.

Zohouri, M., Sabzali, M., & Golmohammadi, A. (2024). Ethical considerations of ChatGPT-assisted article writing. Synesis (ISSN 1984-6754), 16(1), 94-113.

Lijing Gao is an Assistant Professor of Agricultural Science Communication at the University of Missouri, USA. Her research examines science and risk communication, with a focus on how language, cultural contexts, and social values shape public meaning-making, attitudes, and behavioral responses to emerging technologies. Her current work explores public perceptions and discourse surrounding controversial technology. She has published in leading journals, including International Journal of Communication, Agriculture and Human Values, Precision Agriculture, and Rural Sociology.

Ruanjia Liu is an Assistant Professor of Practice in Accounting at Moravian University, US. Her research focuses on governmental accounting, not-for-profit accounting, and accounting education. Her recent work focuses on using natural language processing to analyze the 10-K filings of U.S. firms. Her job market paper was published in the peer-reviewed Journal of Information Systems (JIS).