Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-05484-6 OPEN

Writing without borders: AI and cross-cultural convergence in academic writing quality

Arjun Prakash [1✉], Shruti Aggarwal [1], Jeevan John Varghese [1] & Joel John Varghese [2]

English has become the dominant language in global academic publishing, facilitating cross-border collaboration while reinforcing structural barriers for non-native English-speaking researchers. This study examines the evolution of academic writing quality in social sciences abstracts from 2012 to 2024, focusing on disparities across linguistic, regional, economic and gender-based classifications. Using over one million English-language abstracts retrieved from the Web of Science, the study evaluates writing complexity through readability metrics such as the Flesch-Kincaid Grade Level. A mixed generalised linear model (GLM) is employed to identify key factors influencing writing quality, with particular emphasis on internet access. To assess the potential impact of large language models (LLMs) such as ChatGPT, the analysis incorporates a lexical tracking approach that monitors the frequency of adjectives and adverbs commonly associated with AI-generated content. The findings reveal a global improvement in writing complexity, with non-native English-speaking countries showing notable advances. China, initially lagging in English writing standards, has surpassed traditional leaders such as the United States, signalling a shift in global academic communication. Enhanced digital infrastructure and the adoption of AI-assisted writing tools appear to play a contributory role in this convergence. These results offer empirical insights into how technological advancements are reshaping scholarly expression and mitigating long-standing linguistic and structural disparities. The study provides evidence-based guidance for policy-makers, educators and research institutions seeking to enhance the accessibility, inclusivity and quality of academic writing across diverse global contexts.

[1] Department of Humanities and Social Sciences, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India. [2] School of Engineering, Cochin University of Science and Technology Kerala, Kochi, India. ✉email: psarun.mangad@gmail.com

Introduction

Academic writing remains a central gatekeeping mechanism in global scholarship (Curry & Lillis, 2018). While it enables the circulation of knowledge, it also entrenches linguistic and regional hierarchies (Hanauer, Sheridan & Englander, 2019). Scholars from non-native English-speaking and lower-resource contexts often face barriers in meeting both the substantive expectations for publication and the dominant norms of linguistic clarity, coherence and style (Flowerdew, 2008; Luo & Hyland, 2016; McKinley & Rose, 2018). These challenges are intensified by structural inequalities in global academia that disadvantage scholars outside core institutional networks (Bennett, 2014; Murray, 2015; Zeng & Yang, 2024).

While the global dominance of English has facilitated cross-border collaboration (Dearden, 2014) and expanded educational access through English as a Medium of Instruction (Galloway, 2020; Tong et al., 2020), it continues to sustain inequalities in global academic publishing. This hegemony, shaped by historical patterns of linguistic imperialism (Zeng & Yang, 2024), places additional burdens on scholars from non-native English-speaking contexts, who must navigate complex disciplinary content while meeting unfamiliar linguistic expectations (Clark & Yu, 2021; Zhang & Hasim, 2023). These constraints contribute to their continued underrepresentation in high-impact journals and broader academic discourse (McKinley & Rose, 2018).

These disparities are embedded in a broader core–periphery dynamic where North America and Western Europe dominate knowledge production, marginalising contributions from semi-peripheral and lower-income regions (Mosbah-Natanson, Gingras, 2014; Wagner et al., 2001). Scholars from these regions frequently report that their work is perceived as local or descriptive rather than theoretical or globally relevant (Bennett, 2014), a perception shaped partly by linguistic hierarchies and publishing standards favouring native speakers (McKinley & Rose, 2018; Murray, 2015). In addition to these structural barriers, differences in writing styles and cognitive-cultural framing further challenge non-native speakers' ability to engage with dominant academic norms (Wang & Chen, 2013).

Digital infrastructure, particularly internet access, plays an increasingly critical role in shaping global academic participation. Improved connectivity expands access to scholarly materials and digital learning tools, reducing barriers faced by researchers in underrepresented regions. Studies have shown that internet availability enhances academic confidence), supports research productivity (Xu & Reed, 2021) and facilitates more effective engagement with online learning platforms (Huang et al., 2022), even as structural challenges remain in regions like Africa (Oyelaran-Oyeyinka & Adeya, 2004).

Artificial Intelligence (AI), described as one of the most transformative innovations since the Palaeolithic era and even referred to as the "new oil" (Holmes & Tuomi, 2022), has significantly transformed academic writing practices. Tools such as Grammarly provide real-time grammar and style feedback, enabling non-native speakers to better meet publication standards. More recently, large language models (LLMs) such as ChatGPT have emerged as powerful tools capable of generating fluent, coherent text and assisting in manuscript preparation (Akhtom et al., 2023; Gao et al., 2023). These technologies offer the potential to level the linguistic playing field, but they also raise concerns about transparency and epistemic homogenisation (Cheng, Calhoun & Reedy, 2025; De Maio et al., 2024).

Emerging studies have highlighted the risks of overreliance on AI-generated text. While AI can improve surface-level clarity, it may struggle with originality, creativity and alignment with journal-specific requirements (De Maio et al., 2024; Gao et al., 2023). Concerns about reduced linguistic and epistemic diversity in AI-assisted manuscripts are growing, along with calls for clearer disclosure policies and author accountability (Hosseini et al., 2023; Nature Machine Intelligence, 2022). Identifying AI-generated content remains a challenge (Clark et al., 2021), and recent studies stress the need for robust detection and ethical guidelines (Liebrenz et al., 2023; Nature Editorial, 2023). Despite these concerns, AI tools continue to offer meaningful benefits, especially to non-native English speakers, by improving writing quality and expanding access to academic discourse (Alharbi, 2023; Warschauer et al., 2023).

The global implications of English-language dominance and the growing integration of AI tools in academic writing have garnered increasing scholarly attention. However, few empirical studies have systematically examined large-scale trends in writing quality or assessed how digital infrastructure and AI-assisted tools are reshaping scholarly expression across diverse contexts. This study addresses that gap by analysing over one million social sciences abstracts published between 2012 and 2024. It employs established readability metrics to evaluate writing complexity, investigates the influence of digital access and examines whether recent advances, particularly large language models like ChatGPT, are narrowing these divides or entrenching them further. In doing so, it interrogates the notion of a 'borderless' academic writing landscape in an era shaped by global technological convergence.

Research questions. This study seeks to answer the following questions:

a. How has the academic writing quality of social sciences abstracts evolved from 2012 to 2024, and how does it vary between native and non-native English-speaking countries, and across gender, regional and income-based classifications?
b. What are the key factors influencing writing quality in academic abstracts, and how does internet access contribute to these outcomes?
c. To what extent has the adoption of large language models (LLMs) such as ChatGPT influenced language usage trends and contributed to any observed improvements in writing quality?

Materials and methods

Data and variables. This section outlines the dataset, variable construction and methodological framework used to analyse academic writing quality and the influence of AI and digital access.

The study utilises a dataset of academic abstracts obtained from the Web of Science Index, restricted to entries listed under the Social Sciences Citation Index (SSCI). The dataset, retrieved on March 15, 2024, comprises English-language research articles published between 2012 and 2024 by three major academic publishers. To ensure consistency and relevance, the following inclusion criteria were applied: document type limited to "Article," language designated as "English" and indexing under the SSCI. Records were excluded if they were duplicates, lacked abstracts or originated from publishers outside the three selected publishing houses. The final dataset consisted of approximately 1.03 million articles deemed suitable for analysis.

The data were organised and cleaned using Microsoft Excel, with each row representing a unique article. Metadata fields include author names, corresponding author affiliation and country, publisher information, open access status, funding acknowledgements and abstract text. While the abstracts served as the primary material for evaluating academic writing quality, measured through established readability metrics, the accompanying metadata enabled analysis of variations in writing complexity and potential AI influence across dimensions such as author gender, geographic origin, funding status and other publication characteristics.

The dependent variable in this study is the quality of academic writing, measured through readability scores calculated for each abstract using a Python-based text analysis framework. The independent variables encompass a range of publication characteristics derived from the article-level metadata. The author's gender was inferred using the Gender API, applying a confidence threshold of 80 per cent for inclusion. The country of the corresponding author was extracted from the affiliation data and subsequently classified by income level (High, Upper-Middle and Lower-Income) and geographic region (East Asia and Pacific; Europe and Central Asia; Latin America and the Caribbean; Middle East and North Africa; North America; South Asia; and Sub-Saharan Africa) based on the World Bank classification. Countries were also categorised as either native or non-native English-speaking, depending on whether English is regarded as a native language. Additional explanatory variables in the analysis include publication year, funding status, open access status and internet penetration, which is measured by the number of fixed broadband subscriptions per 100 individuals based on World Bank data.

To evaluate the potential influence of large language models (LLMs) on academic writing, this study adopts a lexical tracking approach based on Liang et al. (2024). In their analysis, Liang et al. generated a large sample of AI-produced texts using ChatGPT and compared them to human-written academic material. Using a frequency-based method, they calculated the relative overuse of individual words in AI-generated texts and identified the 100 adjectives and 100 adverbs most strongly associated with LLM outputs. These validated lists were adopted in our study as proxy indicators of potential LLM influence. The frequency of each term was measured at the abstract level using Microsoft Excel to assess stylistic shifts across groups classified by language, region, income level and gender. Appendix Table I presents the full set of variables used in the analysis, while Appendix Tables II and III list the keywords drawn from Liang et al. (2024).

Quantitative and econometric analysis. This study employs a quantitative framework combining readability analysis, econometric modelling and AI-related measures to examine the factors influencing academic writing quality. Readability tests evaluate writing complexity based on sentence length, word difficulty and syllable counts. The Flesch-Kincaid Grade Level estimates the educational level required for comprehension, focusing on sentence and syllable structure. The Gunning Fog Index captures cognitive effort by identifying complex words containing three or more syllables, while the SMOG Index assesses reading difficulty by counting polysyllabic words. While the primary results are based on Flesch-Kincaid scores, the Gunning Fog and SMOG indices are also reported in the Appendix to demonstrate consistency and robustness.

Longer sentences and higher syllable counts are generally associated with greater linguistic complexity, which often corresponds to more intricate syntax and increased cognitive processing demands (Flesch, 1979). Readability metrics are therefore useful for capturing structural and lexical elements that contribute to writing sophistication (Vajjala & Meurers, 2012). Prior studies by Beers and Nagy (2009) and Bi and Jiang (2020) further show that syntactic complexity, including the use of longer clauses and diverse sentence structures, contributes to perceived writing quality. However, readability metrics also highlight a potential trade-off: increased linguistic complexity may enhance the perceived sophistication of writing but simultaneously reduce accessibility, particularly for non-native English-speaking audiences.

The analysis proceeds in three stages, each addressing a distinct aspect of writing quality. First, it evaluates changes in readability over time and across key demographic categories, including author gender, national income level, regional grouping and leading publishing countries (USA, UK and China). Graphical representation is used to highlight differences and emerging patterns across these dimensions.

Second, the study examines the determinants of writing quality using a mixed generalised linear model (GLM). This modelling strategy accommodates the nested nature of the dataset, where individual publications are grouped within countries with differing social, economic and institutional contexts. Fixed effects are used to estimate the influence of variables such as internet access, open access status and author gender, while random effects account for unobserved variation at the country level. Allowing slopes to vary for internet penetration further reveals how the relationship between digital infrastructure and writing quality may differ across countries. The Annexure provides a detailed explanation of the model formulation and interpretation of fixed and random components.

Third, the study investigates how large language models (LLMs) may shape academic writing patterns. It calculates the frequency of 100 adjectives and 100 adverbs identified by Liang et al. (2024) as more commonly used in AI-generated texts and compares their distribution across demographic and regional categories. Graphical methods illustrate how the presence of these terms has changed over time, providing insight into the potential influence of LLMs on linguistic practices in academic abstracts.

Results

This section presents the empirical findings on writing complexity, examining variation across demographic and linguistic groups, the influence of key publication-related factors and patterns potentially associated with LLM usage.

Geographic, economic and gender disparities in writing quality. This section provides the variation of the Flesch-Kincaid Reading Grade across different demographic and economic groups. The analysis of other readability tests for these groups is provided in the Appendix for a robustness check (Appendix Figure II to Figure XV).

Figure 1 displays the average readability scores by various demographic and economic factors. As expected, countries where English is the native language demonstrate higher readability scores, reflecting more complex sentence structures that require advanced reading comprehension. Similarly, publications from high-income countries generally exhibit higher readability scores compared to those from lower-income countries, suggesting differences in academic writing standards. Gender-based differences are also evident. On average, female first-authored abstracts exhibit higher readability scores than male-authored ones, indicating potential differences in the complexity of written content. Figure 2 presents regional patterns, where South Asia records the lowest average scores, while East Asia & Pacific and Sub-Saharan Africa show higher scores, indicating substantial regional variation in writing complexity.

The trends in Flesch-Kincaid Grade Scores from 2012 to 2024 across various categories highlight notable trends in the evolution of academic writing quality. Figure 3 depicts a general upward trend with occasional spikes in readability scores for the study

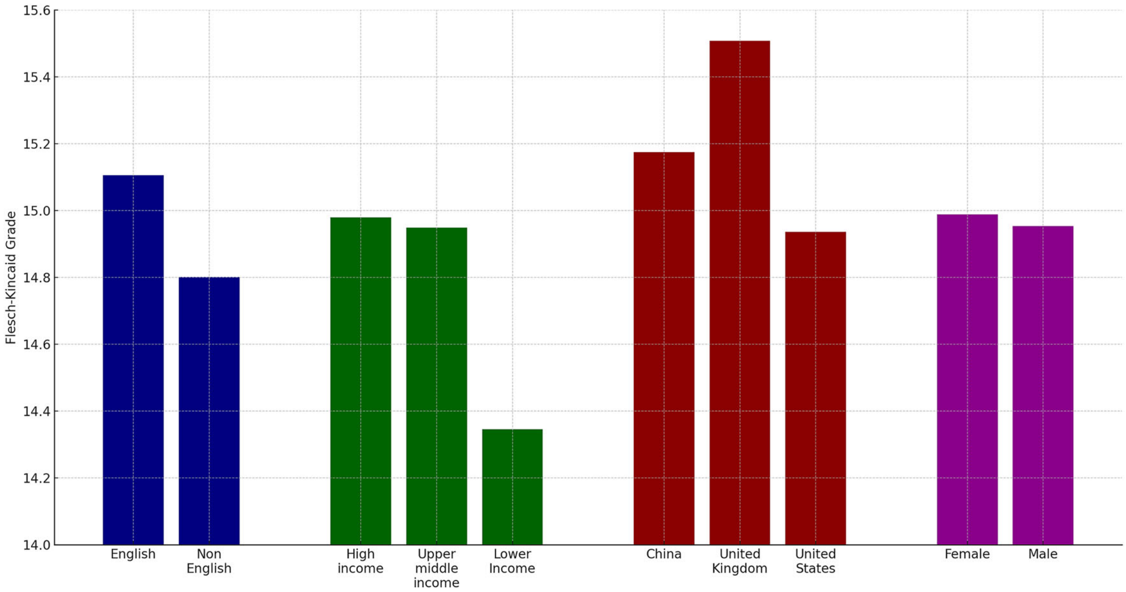

Fig. 1 Average Flesch-Kincaid Grade by demographic and economic factors. This figure displays the average Flesch-Kincaid Grade level of academic writing quality across various demographic and economic categories, including language (English vs. Non-English), income level (High Income, Upper Middle Income, Lower Income), country (China, United Kingdom, United States), and gender (Female, Male).

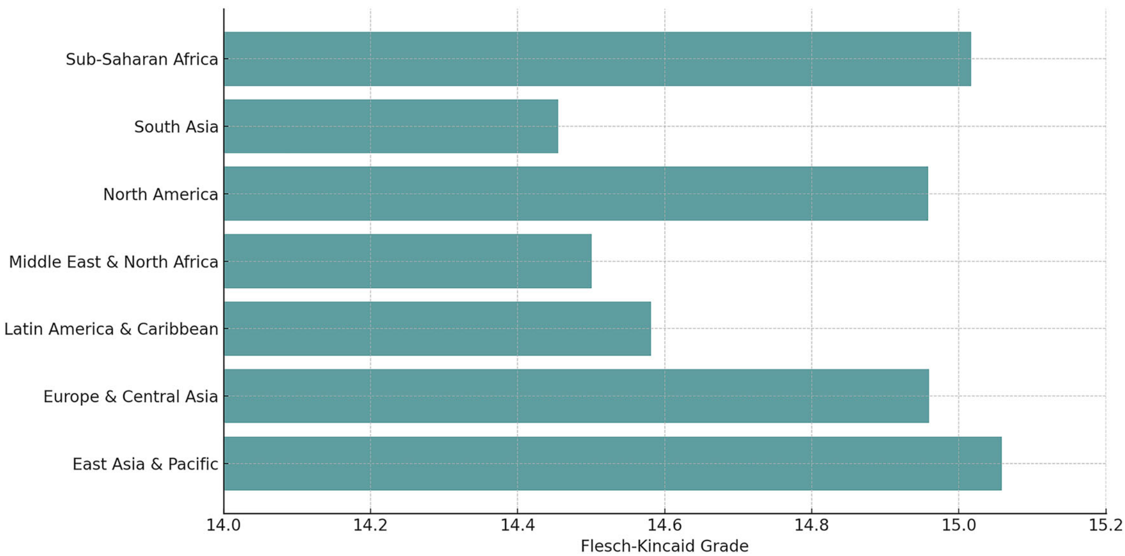

Fig. 2 Average Flesch-Kincaid Grade by region. This figure presents the average Flesch-Kincaid Grade level across different global regions, highlighting regional disparities in academic writing quality. Regions include Sub-Saharan Africa, South Asia, North America, Middle East & North Africa, Latin America & Caribbean, Europe & Central Asia, and East Asia & Pacific.

period, indicating improved writing quality. However, a notable dip occurred in 2020, likely impacted by the COVID-19 pandemic. Figure 4 provides a comparative analysis of the evolution of writing quality between English-native and non-English-native countries. While publications from the English-native countries initially exhibited a higher reading score, the non-English speaking countries have made significant improvements, closing the gap. This convergence suggests a levelling of global academic writing standards.

Figure 5 depicts the temporal progression of writing complexity across income groups. Although high-income countries maintained the highest scores for much of the study period, upper-middle-income countries eventually surpassed them. Figure 6 compares the top three publishing countries—China, the United Kingdom and the United States. The United Kingdom consistently recorded the highest readability scores, reflecting greater linguistic complexity. Notably, China, which began with the lowest scores in 2012, demonstrated significant improvement over time, overtaking the United States by 2019 and approaching the United Kingdom by 2024. This trajectory highlights China's rapid progress in English academic writing and challenges the conventional dominance of native English-speaking countries. Appendix Figure II presents gender-based trends in readability scores. Although abstracts with female first authors initially demonstrated higher scores, the gap narrowed after 2020, and male-authored abstracts eventually surpassed them, suggesting a shift in English writing proficiency across genders within academic contexts.

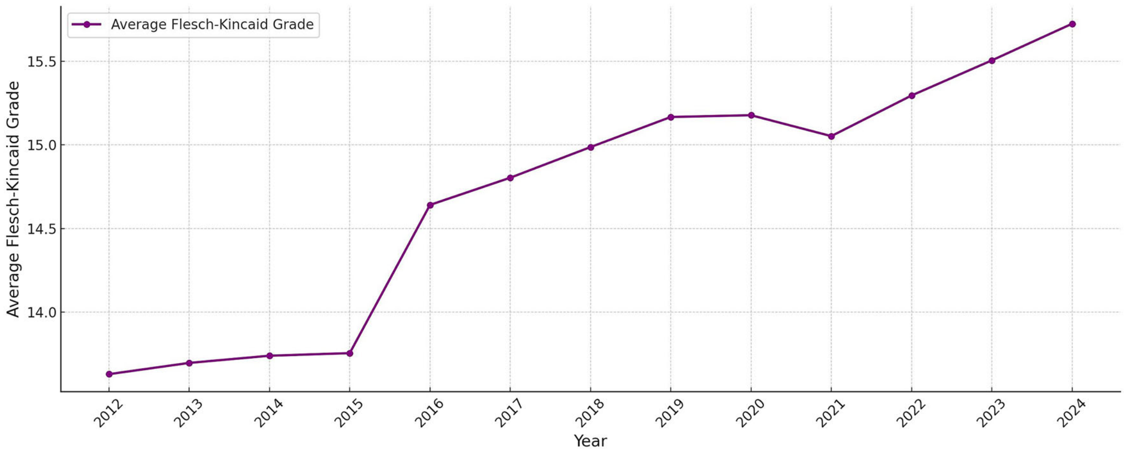

Fig. 3 Average Flesch-Kincaid Reading Score over the years. This figure illustrates the trend in average Flesch-Kincaid Grade levels for academic abstracts from 2012 to 2024, showing how writing complexity has evolved over time.

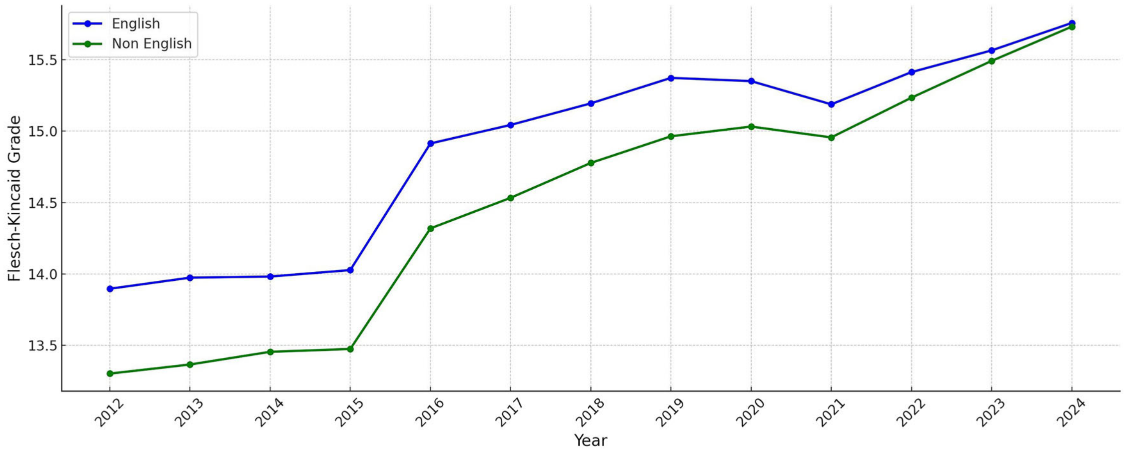

Fig. 4 Average Flesch-Kincaid Reading for native English-speaking and non-native English-speaking countries. This figure compares the average Flesch-Kincaid Grade levels between native English-speaking and non-native English-speaking countries from 2012 to 2024, demonstrating trends in writing quality across linguistic backgrounds.

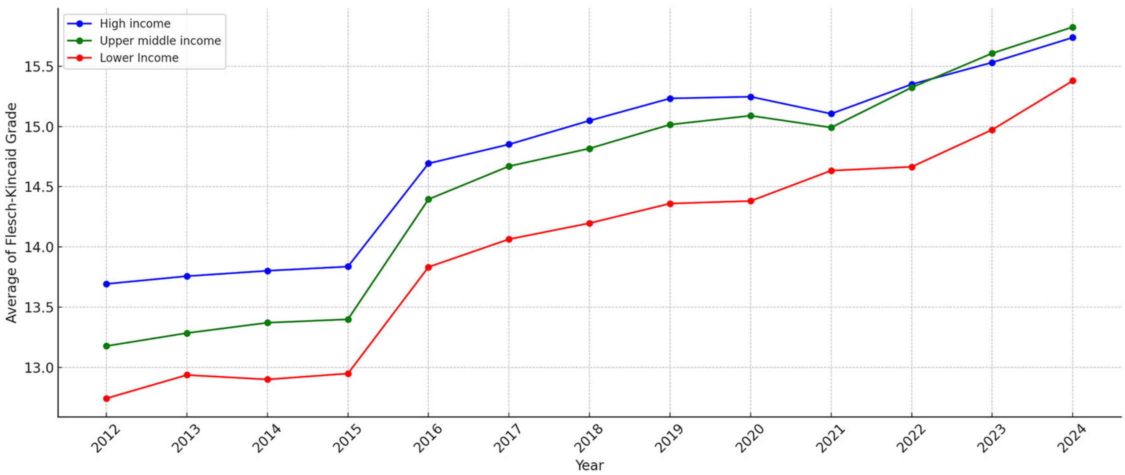

Fig. 5 Average Flesch-Kincaid Reading for income groups. This figure shows the average Flesch-Kincaid Grade levels for academic writing across different income groups (High Income, Upper Middle Income, Lower Income) over the period from 2012 to 2024.

Econometric analysis of factors influencing writing quality. Table 1 presents the results of a mixed-effects generalised linear model (GLM) analysing the factors associated with academic writing quality, as measured by Flesch-Kincaid Grade Scores. Internet penetration, measured by fixed broadband subscriptions per 100 people, emerged as a significant predictor and was positively correlated with writing complexity. This suggests that greater internet access is associated with the production of more complex academic texts.

Native language also showed a significant effect. Publications from native English-speaking countries generally exhibited higher

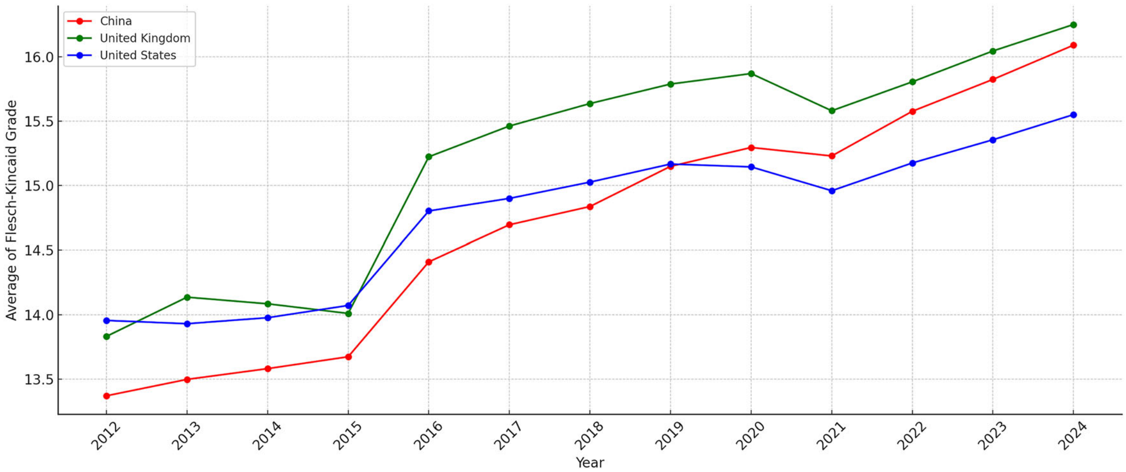

Fig. 6 Average Flesch-Kincaid Reading among the top 3 publishing countries. This figure displays the average Flesch-Kincaid Grade levels for the top three publishing countries—China, the United Kingdom, and the United States—over the period from 2012 to 2024.

Table 1 Mixed GLM-factors impacting academic writing.

Flesch Kincaid Grade	Coefficient	Std. err.	*P* > *z*
Internet penetration	0.0315	0.0044	0.000
Native Language English	0.3280	0.1231	0.008
Native Language Non-English (omitted)			
Female	0.0311	0.0103	0.003
Male (omitted)			
Funded	−0.2972	0.0107	0.000
Non-Funded (omitted)			
Open Access	−0.1969	0.0115	0.000
Non-Open Access (omitted)			
High Income	0.2662	0.2573	0.301
Upper Middle Income	0.5927	0.2380	0.013
Lower Middle Income	0.6351	0.2255	0.005
Low Income (omitted)			
East Asia Pacific	0.2205	0.1448	0.128
Europe and Central Asia	0.2069	0.1391	0.137
Latin America and Caribbean	0.5089	0.1634	0.002
South Asia	0.3648	0.1793	0.042
North America	−0.0149	0.2306	0.949
Sub-Saharan Africa	0.9702	0.1669	0.000
Middle East and North Africa (omitted)			
_cons	13.2452	0.2484	0.000

This table presents the results of a Mixed Generalised Linear Model (GLM) assessing the impact of various demographic, economic, and academic factors on the quality of academic writing, as measured by the Flesch-Kincaid Grade level. Coefficients indicate the relationship strength and direction between each factor and the Flesch-Kincaid Grade. Standard errors (std. err) and *p*-values ($P > z$) are included to indicate the precision and statistical significance of these relationships. Factors considered include internet penetration, native language, gender, funding status, open access status, income levels, and geographical regions. The omitted categories for comparison are provided in parentheses.

levels of writing complexity compared to those from non-native English-speaking countries. Additionally, the gender of the first author was significantly associated with writing quality, with female-authored publications displaying higher readability scores than their male-authored counterparts. Funding status and open-access publication were also found to be influential. Notably, publications that were funded or published open access tended to have lower readability scores than those that were unfunded and not open access.

The impact of income and regional groups on writing quality became modest with controls for internet penetration and native language. High-income status did not have a statistically significant impact relative to low-income status, while upper-middle- and lower-middle-income classifications were positively and significantly associated with writing quality. Regionally, publications originating from Latin America and the Caribbean, as well as Sub-Saharan Africa, demonstrated significantly higher readability scores compared to the reference group of the Middle East and North Africa. In contrast, no significant differences were observed for publications from North America, Europe and Central Asia or East Asia and the Pacific.

Influence of AI on academic writing. This section analyses the evolving use of linguistic elements commonly associated with large language models (LLMs) such as ChatGPT across different demographic categories. Figure 7 presents longitudinal trends in the usage of specific adjectives and adverbs frequently over-represented in AI-generated texts, comparing native English-speaking and non-native English-speaking countries. Initially, native English-speaking countries exhibited higher usage of these terms. However, from 2021 onwards, non-native English-speaking countries demonstrated a notable increase, ultimately surpassing their native counterparts. This shift suggests a wider global diffusion of AI-assisted writing tools. Further differences emerge across economic classifications. Upper-middle-income and lower-income countries recorded higher frequencies of these stylistic markers compared to high-income countries, indicating broader adoption of AI-supported linguistic conventions across income strata (Appendix Figures III and IV).

Figure 8 illustrates trends in the usage of unique adjectives and adverbs commonly associated with large language models (LLMs) across the top three publishing countries—China, the United Kingdom and the United States. The figure reveals a marked increase in the use of LLM-associated linguistic patterns, particularly in China, which shows a significant surge post-2021. This rise corresponds with the previously observed improvement in China's overall writing complexity, reinforcing a potential link between increased LLM usage and enhanced readability scores.

Figure 9 and Appendix Figure V provide insights into the adoption of specific terms identified as being commonly overused by LLMs, such as meticulous and intricate. Between 2022 and 2024, the term *meticulous* experienced a 7.8-fold average increase in usage across all groups. Non-native English-speaking authors exhibited a tenfold increase compared to a 2.28-fold increase among native English speakers. Gender-based patterns were also evident. Male first-authored studies used the term 13 times more

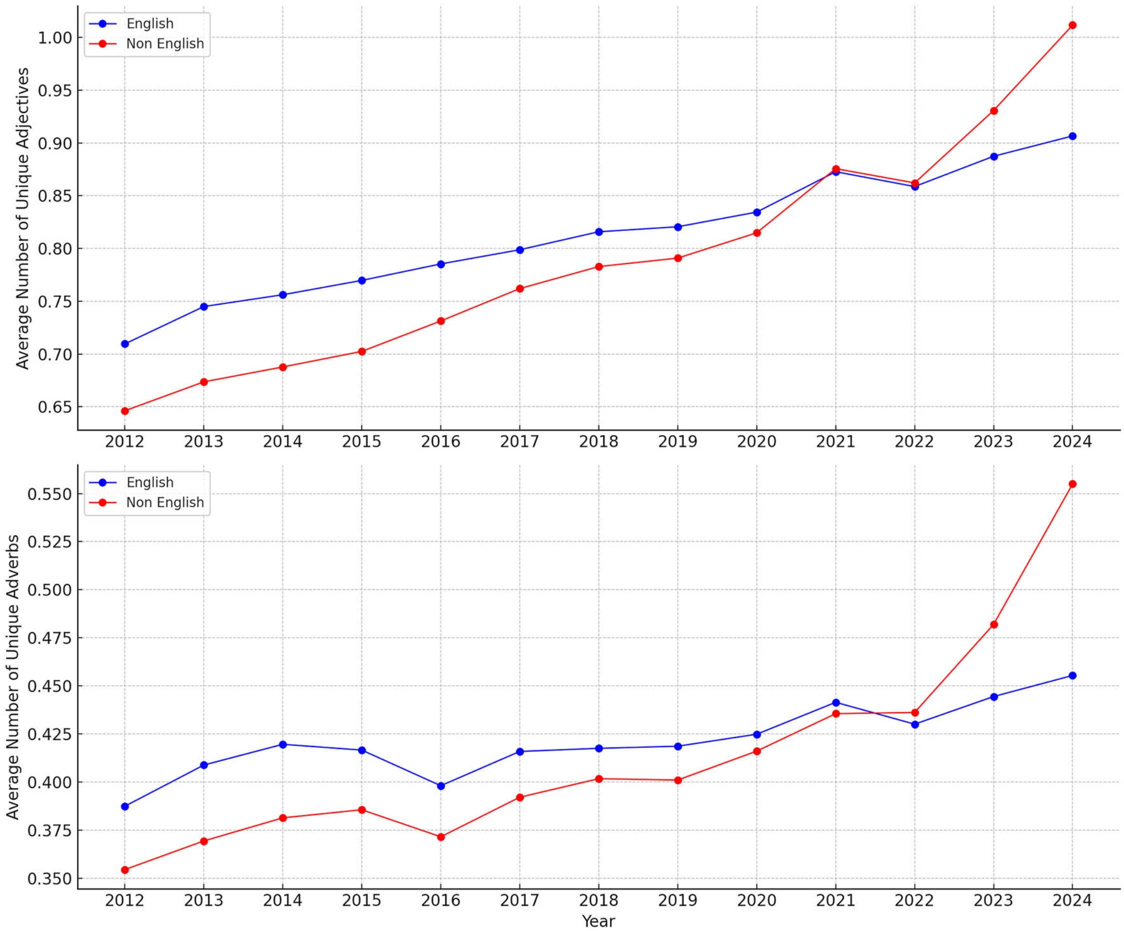

Fig. 7 Trends in the use of unique adjectives and adverbs in academic writing: English vs. non-English native countries. This figure shows the trend in the usage of unique adjectives (subplot 1) and unique adverbs (subplot 2) commonly associated with AI-generated content in academic writing across English-speaking and non-English-speaking countries from 2012 to 2024.

while female studies showed a fourfold increase, suggesting notable gender-based adoption patterns. Regionally, the term's usage increased by 1.8 times in the United States, four times in the United Kingdom and an impressive seventeen times in China. These patterns highlight China's substantial adoption of LLM-influenced linguistic tools compared to other regions.

Similarly, Appendix Figure V shows that the term *intricate* saw a 5.9-fold increase overall. Non-native English speakers increased their usage eightfold while native English speakers showed a 2.8-fold increase. Among high-income countries, there was a 4.6-fold increase in usage, with China and the UK recording significant rises of 10.18 and 6.89 times, respectively. In contrast, usage of *intricate* declined slightly in the United States.

Discussion

This section interprets the study's findings in relation to existing debates on writing quality, digital access and AI-driven language support in academic publishing.

The study examined how the complexity of academic writing in social sciences abstracts evolved from 2012 to 2024 across geographic, linguistic and economic contexts. It further explored the influence of digital infrastructure and large language models (LLMs) on these trends, using large-scale readability metrics and mixed-effects econometric modelling.

The findings reveal a steady improvement in writing complexity globally, with particularly notable gains among authors from non-native English-speaking and lower-income countries. This convergence suggests that digital tools and institutional emphasis on communication standards may be narrowing historical disparities in scholarly expression. China's rise in readability scores, now surpassing some native English-speaking countries, mirrors earlier evidence that national policy initiatives have enhanced English proficiency and research visibility (Gao & Zheng, 2019).

Regression analysis underscores broadband access as a key predictor of writing quality. This is consistent with previous work linking digital connectivity to increased research output and academic capacity (Xu & Reed, 2021). Tools like Grammarly and other AI-assisted platforms are known to help researchers refine grammar and style, particularly in contexts where access to professional editing services is limited (Katsnelson, 2022; Ghufron & Rosyida, 2018). von Garrel, Mayer (2023) further show that such tools are widely used among university students, including in the social sciences, for tasks including translation, content creation and clarity improvement. Nonetheless, their limitations in assessing argumentative clarity and disciplinary depth have also been noted in recent empirical studies (Al-Kadi, 2025).

A notable pattern in our results is the increased use of adjectives and adverbs typically associated with LLMs, especially in abstracts from non-native and lower-income regions. This suggests a growing reliance on AI tools in academic writing (Cui, 2025). Similar large-scale trends have been reported by Geng and Trotta (2024), who examined over one million arXiv abstracts and found a marked rise in ChatGPT-associated word frequencies following its release, particularly in computer science. Their

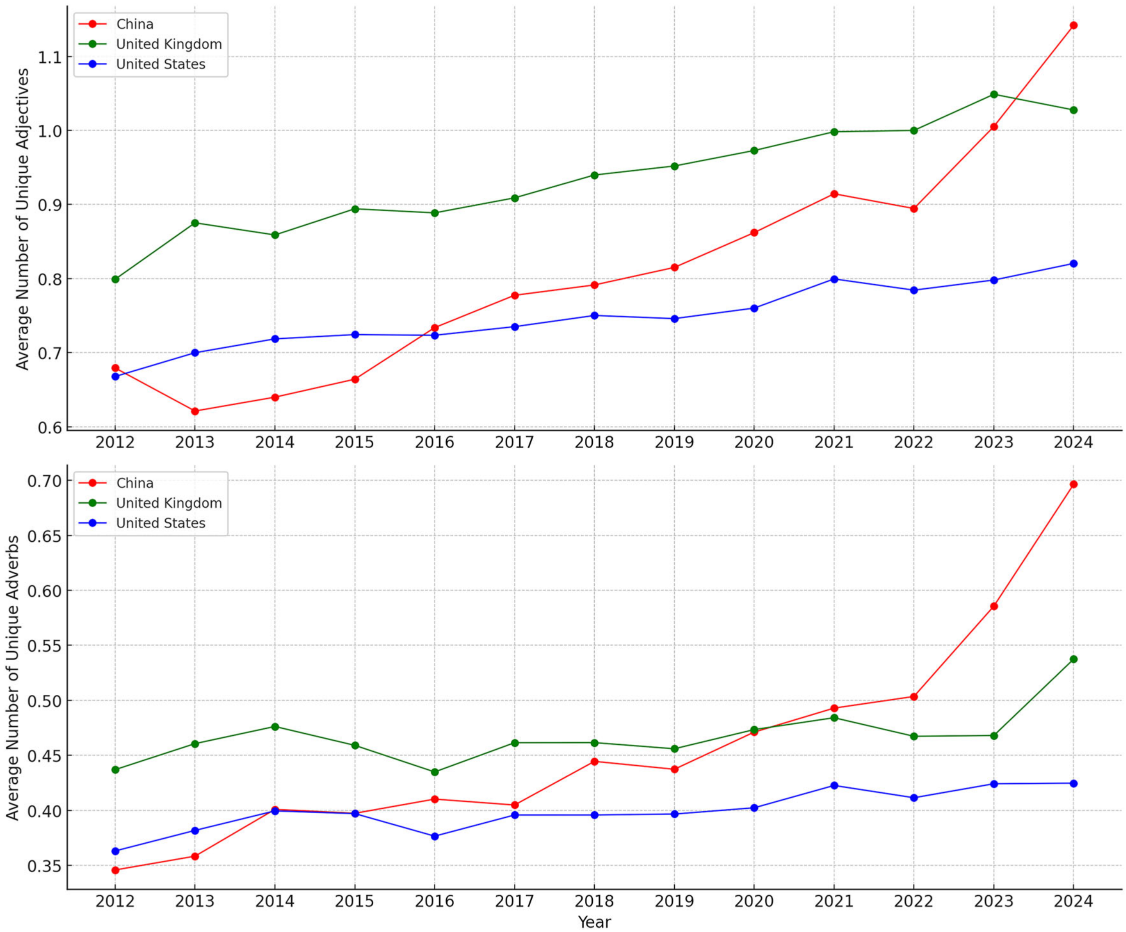

Fig. 8 Evolution of unique adjective and adverb usage by the top 3 publishing countries. This figure presents the evolution of unique adjectives (sub-plot 1) and unique adverbs (sub-plot 2) commonly associated with AI-generated content in academic writing in the top three publishing countries—China, the United Kingdom, and the United States—from 2012 to 2024.

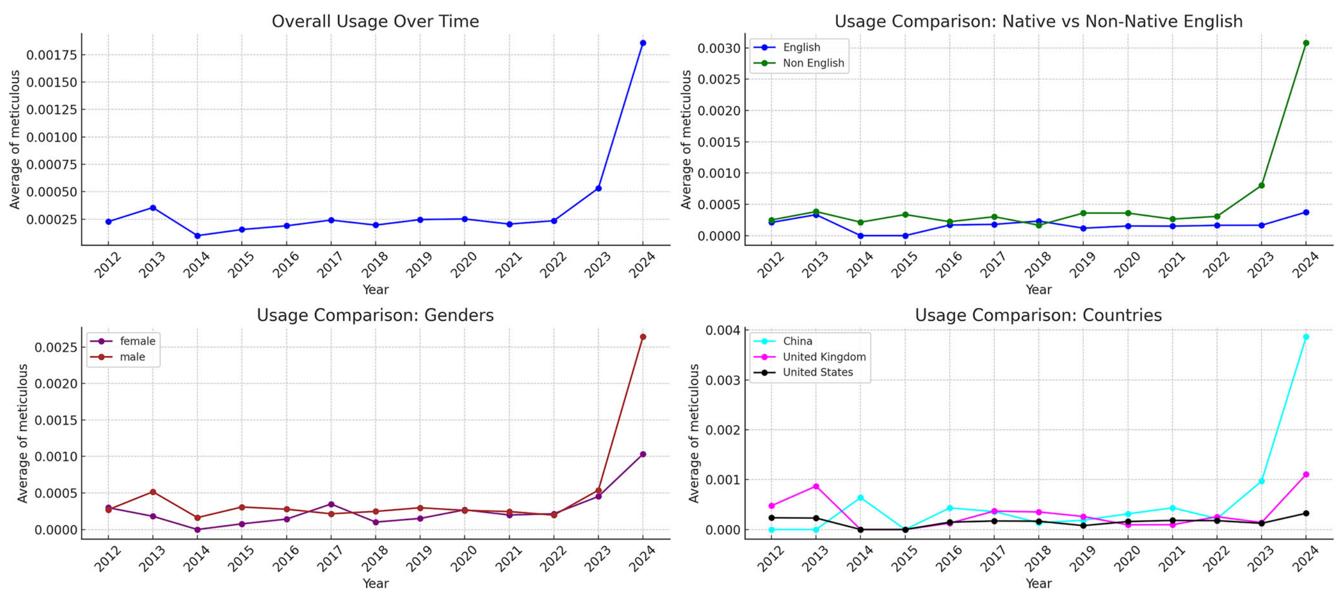

Fig. 9 Trend in usage of the term "meticulous". This figure tracks the usage frequency of the word "meticulous," identified as characteristic of AI-generated text, across various dimensions. The figure includes an overall trend line showing how the term's usage has changed over time from 2012 to 2024. Additionally, it compares the usage between native and non-native English-speaking countries, gender groups (male vs. female), and among the top three publishing countries (China, the United Kingdom, and the United States).

findings offer independent confirmation that LLMs are reshaping writing styles at scale. Consistent with this, our study observes increased use of LLM-associated linguistic features across social sciences abstracts, reflecting broader stylistic shifts in academic writing. Recent studies further affirm that LLMs enhance structural and syntactic fluency, particularly for non-native English speakers (Li et al., 2024) and support writing productivity by assisting with editing, formatting and summarising tasks (Korinek, 2023). However, these benefits are not universal. For instance, Bašić et al. (2023) found that access to ChatGPT did not significantly improve student essay performance, suggesting that effectiveness may depend on user familiarity and context.

While our findings highlight the growing presence of AI-influenced language in academic abstracts, they also point to broader implications for scholarly communication. Emerging literature suggests that while such tools can enhance standardisation and linguistic accessibility, they may also contribute to stylistic convergence at the expense of narrative depth and originality. Studies have shown that AI-generated texts, though grammatically polished, often exhibit predictable structures and limited creative nuance (Conroy, 2023; Kong & Liu, 2024), potentially diluting the richness of academic discourse. In addition, the increasing indistinguishability of AI-generated content (Clark et al., 2021) complicates efforts to uphold transparency and authorship accountability. Although models like ChatGPT can produce human-like text (Gao et al., 2023), they frequently struggle with context-specific reasoning, disciplinary conventions and formatting requirements (Oates, Johnson, 2025). Furthermore, the reliance on training data skewed toward dominant linguistic and cultural norms raises concerns about embedded bias and the reinforcement of existing academic hierarchies (Humble & Mozelius, 2022). These risks are further compounded by the commercialisation of advanced AI platforms, which may limit equitable access for scholars in less-resourced settings (Liebrenz et al., 2023; Michalak & Ellixson, 2025).

To address these concerns, journals and academic institutions should implement robust AI-detection mechanisms such as watermarking, metadata tracking and algorithmic screening, and collaborate with AI developers to refine these tools in line with evolving capabilities. An essential first step is mandating the disclosure of AI use at the time of submission, with guidelines that clearly distinguish between AI-generated content, produced autonomously by generative tools like ChatGPT, and AI-enabled content, where tools are used solely to enhance grammar or style without contributing substantive content. These guidelines should be regularly updated to reflect advances in AI and ensure ongoing ethical compliance. In parallel, institutions and research funders must invest in capacity-building initiatives that equip scholars with the knowledge to engage with these technologies responsibly. Expanding access to open-source AI tools and promoting the use of linguistically and culturally diverse training datasets can help mitigate systemic bias. Transparent disclosure of AI involvement in data handling, model use and content generation is critical to maintaining trust in AI-assisted research and upholding scholarly accountability.

Overall, this study provides large-scale empirical evidence of an overall improvement in academic writing quality, particularly in regions historically marginalised due to linguistic and infrastructural disadvantages. These gains appear to be supported by expanded internet access and the growing integration of AI tools in academic writing. However, realising the full potential of these technologies will depend on how inclusively and ethically they are implemented. Technological progress should enhance rather than compromise the plurality and scholarly rigour of global academic communication.

Conclusion

This section synthesises the study's key findings in relation to the research questions and discusses their broader implications for linguistic equity and scholarly participation.

In response to the first research question, the study found a consistent upward trend in writing complexity from 2012 to 2024, with especially notable gains among authors from non-native English-speaking countries. For the second question, internet access—measured by fixed broadband subscriptions—emerged as a significant predictor of writing quality; countries with stronger digital connectivity tended to produce more syntactically complex texts, suggesting that digital infrastructure may enable more advanced academic expression. Addressing the third question, lexical analysis showed that LLM-associated language patterns have become increasingly common, particularly in abstracts by non-native English-speaking and lower-income authors. This rise coincided with the period of most notable writing improvements for these groups. Although the method does not establish causality, the alignment between linguistic shifts and quality gains suggests that generative tools may be contributing to changes in academic expression.

Historically, academic publishers have offered optional language editing services to support authors facing linguistic barriers, but such services are often financially inaccessible in under-resourced settings. In contrast, LLM-enabled tools offer a more scalable and affordable alternative, allowing researchers to focus on content rather than surface-level correction. As these tools become more widely adopted, they may support broader inclusion in academic publishing.

This study does not advocate replacing human authorship. While AI tools can enhance clarity and style, the foundation of academic writing remains rooted in original thought, disciplinary expertise and critical analysis. Ensuring ethical and transparent use of these technologies is essential for protecting scholarly integrity, intellectual diversity and equitable participation in global research.

Taken together, the findings show that writing quality in social sciences abstracts has improved over time, with the most marked gains among non-native English-speaking and lower-income authors. These shifts appear linked to both structural factors, such as expanded internet access, and the increasing use of AI-based tools like LLMs. The lexical patterns observed suggest that such tools may be helping scholars achieve greater fluency and coherence, especially where access to formal editing support is limited. While not a substitute for scholarly judgement, generative tools may help reduce persistent linguistic and infrastructural barriers, contributing to more inclusive global academic communication.

Limitations and future directions

This section outlines the study's main limitations and proposes areas for future research to address unresolved questions and expand the scope of analysis.

First, the dataset is restricted to the Web of Science—Social Sciences Citation Index (SSCI) and includes only articles from three major academic publishers. While this ensures consistency in metadata and classification, it limits the breadth and generalisability of the findings relative to broader databases such as Scopus or Lens.org.

Second, the analysis focuses on abstracts rather than full manuscripts. Although abstracts are widely used and standardised for academic summaries, they do not capture the depth of argumentation, structural organisation or conceptual framing present in complete articles. Additionally, the use of readability metrics offers insight into surface-level linguistic complexity but does not assess deeper qualities such as reasoning clarity or rhetorical flow. Future studies could address this by incorporating qualitative assessments to provide a more comprehensive view of writing quality.

Third, the study uses a fixed set of 100 adjectives and 100 adverbs previously identified by Liang et al. (2024) as stylistically prominent in AI-generated text. These were employed as proxy indicators to explore the potential influence of large language models (LLMs) on academic writing. While this approach provides a useful starting point, the limited scope of the lexical list may not fully capture the broader spectrum of AI-influenced language. We did not independently derive or validate a new lexicon as this was beyond the scope of the present study. Future studies could build on this by conducting original analyses to identify evolving linguistic markers of AI-generated content across disciplines and genres. Additionally, the econometric models used in this study incorporate data only up to 2021, potentially missing more recent shifts in writing practices following the wider adoption of AI tools in academic workflows after the COVID-19 pandemic.

Fourth, gender classification was conducted using the Gender API, applying an 80% confidence threshold. While suitable for large-scale inference, this method may misclassify culturally ambiguous or less common names. Accordingly, conclusions related to gender should be interpreted with appropriate caution.

Future research could further examine the broader influence of artificial intelligence on education, creativity and communication. At the primary and secondary education levels, the integration of AI tools may impact the development of foundational writing skills and raise concerns regarding academic integrity in student assignments. The ability of AI systems to generate creative outputs, such as poetry and narrative prose, also invites reflection on the evolving nature of human creativity and artistic expression in an increasingly automated environment. Moreover, AI-enabled translation and drafting tools hold potential for enhancing professional communication, enabling non-native speakers to produce official documents more independently and promoting greater inclusivity in administrative and institutional contexts.

The ethical and inclusive integration of AI into academic practice will require sustained attention, responsive policy development and sensitivity to disciplinary and regional contexts. Addressing these issues allows the academic community to leverage AI's advantages to foster equitable participation while safeguarding scholarly diversity and integrity.

Data availability

The dataset and replication code used in this study are openly available on Zenodo at the following https://zenodo.org/records/15755757.

Received: 13 September 2024; Accepted: 1 July 2025;
Published online: 09 July 2025

References

Akhtom D, Alyasiri OM, Allogmani E, Salman AM, Sahib TM (2023) Unlocking ChatGPT's title generation potential: an investigation of synonyms, readability, and introduction-based titles. J Theor Appl Inf Technol 101(22):7435–7443

Al-Kadi A (2025) Fostering a 'Human with AI' approach for evaluating students' writing in English. Stud Linguist Cult FLT 13(1):140–159. https://doi.org/10.46687/VFBZ9792

Alharbi W (2023) AI in the foreign language classroom: a pedagogical overview of automated writing assistance tools. Educ Res Int 2023:4253331. https://doi.org/10.1155/2023/4253331

Bašić Ž, Banovac A, Kružić I, Jerković I (2023) ChatGPT-3.5 as writing assistance in students' essays. Humanit Soc Sci Commun 10:750. https://doi.org/10.1057/s41599-023-02269-7

Beers SF, Nagy WE (2009) Syntactic complexity as a predictor of adolescent writing quality: which measures? Which genre? Read Writ 22(2):185–200. https://doi.org/10.1007/s11145-007-9107-5

Bennett K (ed.). (2014) The semiperiphery of academic writing. Palgrave Macmillan UK. https://doi.org/10.1057/9781137351197

Bi P, Jiang J (2020) Syntactic complexity in assessing young adolescent EFL learners' writings: syntactic elaboration and diversity. System 91:102248. https://doi.org/10.1016/j.system.2020.102248

Cheng A, Calhoun A, Reedy G (2025) Artificial intelligence-assisted academic writing: recommendations for ethical use. Adv Simulat 10(22). https://doi.org/10.1186/s41077-025-00350-6

Clark E, August T, Serrano S, Haduong N, Gururangan S, Smith NA (2021) All that's 'human' is not gold: Evaluating human evaluation of generated text. In: Tetreault J, Burstein J, Leacock C (eds.) Proceedings of the ACL-IJCNLP 2021: 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Association for Computational Linguistics. pp. 7282–7296. https://doi.org/10.18653/v1/2021.acl-long.565

Clark T, Yu G (2021) Beyond the IELTS test: Chinese and Japanese postgraduate UK experiences. Int J Biling Educ Bilingual 24(10):1512–1530. https://doi.org/10.1080/13670050.2020.1829538

Conroy M (2023) Scientists used ChatGPT to generate a whole paper from data. Nature 619:443–444. https://doi.org/10.1038/d41586-023-02218-z

Cui Y (2025) What influences college students using AI for academic writing? A quantitative analysis based on HISAM and TRI theory. Comput Educ Artif Intell 8:100391, https://www.sciencedirect.com/science/article/pii/S2666920X25000311

Curry MJ, Lillis T (2018) Global academic publishing: policies, perspectives and pedagogies. Multilingual Matters

De Maio JL, Kabalaki I, Moshtael S, Tejax MA (2024) AI versus students: a study of the capability of ChatGPT to write Model United Nations position papers. Polit Sci Polit 1–6. https://doi.org/10.1017/S1049096524000799

Dearden J (2014) English as a medium of instruction—a growing global phenomenon. British Council. https://www.teachingenglish.org.uk

Flesch R (1979) How to write plain English: a book for lawyers and consumers. Harper & Row

Flowerdew J (2008) Scholarly writers who use English as an additional language: What can Goffman's "Stigma" tell us? J Engl Acad Purp 7(2):77–86. https://doi.org/10.1016/j.jeap.2008.03.002

Galloway N (ed.). (2020) English in higher education: English medium. Part 1: Literature review. British Council

Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, Pearson AT (2023) Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. npj Digit Med 6(1):75. https://doi.org/10.1038/s41746-023-00819-6. Article

Gao X, Zheng Y (2019) Multilingualism and higher education in Greater China. J Multiling Multicult Dev 40(7):555–561. https://doi.org/10.1080/01434632.2019.1571073

Geng M, Trotta R (2024). Is ChatGPT transforming academics' writing style? Preprint at https://doi.org/10.48550/arXiv.2404.08627

Ghufron MA, Rosyida F (2018) The role of Grammarly in assessing English as a foreign language (EFL) writing. Ling Cult 12(4):395–403. https://doi.org/10.21512/lc.v12i4.4582

Hanauer DI, Sheridan CL, Englander K (2019) Linguistic injustice in the writing of research articles in English as a second language: data from Taiwanese and Mexican researchers. Writ Commun 36(1):136–154. https://doi.org/10.1177/0741088318804821

Holmes W, Tuomi I (2022) State of the art and practice in AI in education. Eur J Educ 57:542–570. https://doi.org/10.1111/ejed.12533

Hosseini M, Rasmussen L, Resnik D (2023) Using AI to write scholarly publications. Account Res 1–9. https://doi.org/10.1080/08989621.2023.2168535

Huang F, Teo T, Scherer R (2022) Investigating the antecedents of university students' perceived ease of using the internet for learning. Interact Learn Environ 30(7):1060–1076. https://doi.org/10.1080/10494820.2019.1710540

Humble N, Mozelius P (2022) The threat, hype, and promise of artificial intelligence in education. Discov Artif Intell 2(1):39. https://doi.org/10.1007/s44163-022-00039-z

Katsnelson A (2022) Poor English skills? There's an AI for that: Machine-learning tools can correct grammar and advise on the style and tone of presentations—but they must be used with caution. Nature 609:208–209. https://doi.org/10.1038/d1586-022-02767-9

Kong X, Liu C (2024) A comparative genre analysis of AI-generated and scholar-written abstracts for English review articles in international journals. J Engl Acad Purp 71:101432. https://doi.org/10.1016/j.jeap.2024.101432. Article

Korinek A (2023) Language models and cognitive automation for economic research (NBER Working Paper No. 30957). National Bureau of Economic Research. http://www.nber.org/papers/w30957

Li J, Huang J, Wu W, Whipple PB (2024) Evaluating the role of ChatGPT in enhancing EFL writing assessments in classroom settings: a preliminary investigation. Humanit Soc Sci Commun 11:1268. https://doi.org/10.1057/s41599-024-03755-2

Liang W, Izzo Z, Zhang Y, Lepp H, Cao H, Zhao X, Chen L, Ye H, Liu S, Huang Z, McFarland DA, Zou JY (2024) Monitoring AI-modified content at scale: a case study on the impact of ChatGPT on AI conference peer reviews. Preprint at. https://doi.org/10.48550/arXiv.2403.07183

Liebrenz M, Schleim S, Hirt J, Bhugra D, Smith A (2023) Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. Lancet Digit Health 5(3):e105–e106. https://doi.org/10.1016/S2589-7500(23)00019-5

Luo N, Hyland K (2016) Chinese academics writing for publication: English teachers as text mediators. J Second Lang Writ 33:43–55. https://doi.org/10.1016/j.jslw.2016.06.005

McKinley J, Rose H (2018) Conceptualizations of language errors, standards, norms and nativeness in English for research publication purposes: an analysis of journal submission guidelines. J Second Lang Writ 42:1–11. https://doi.org/10.1016/j.jslw.2018.07.003

Michalak R, Ellixson D (2025) Fostering ethical AI integration in first-year writing: a case study on human-tool collaboration in artificial intelligence literacy. J Libr Adm 65(3):361–377. https://doi.org/10.1080/01930826.2025.2468136

Mosbah-Natanson S, Gingras Y (2014) The globalisation of social sciences? Evidence from a quantitative analysis of 30 years of production, collaboration and citations in the social sciences (1980–2009). Curr Sociol 62(5):626–646. https://doi.org/10.1177/0011392113498866

Murray N (2015) Standards of English in higher education. Cambridge University Press. https://doi.org/10.1017/CBO9781139507189

Nature Editorial (2023) Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature 613:612. https://doi.org/10.1038/d41586-023-00191-1

Nature Machine Intelligence (2022) Much to discuss in AI ethics. Nat Mach Intell 4(12):1055–1056. https://doi.org/10.1038/s42256-022-00598-x

Oates A, Johnson D (2025) ChatGPT in the classroom: evaluating its role in fostering critical evaluation skills. Int J Artif Intell Educ. https://doi.org/10.1007/s40593-024-00452-8

Oyelaran-Oyeyinka B, Adeya CN (2004) Internet access in Africa: empirical evidence from Kenya and Nigeria. Telemat Inform 21(1):67–81. https://doi.org/10.1016/S0736-5853(03)00023-6

Tong F, Wang Z, Min Y, Tang S (2020) A systematic literature synthesis of 19 years of bilingual education in Chinese higher education: where does the academic discourse stand? Sage Open 10(2). https://doi.org/10.1177/2158244020926510

Vajjala S, Meurers D (2012) On improving the accuracy of readability classification using insights from second language acquisition. In: Tetreault J, Burstein J, Leacock C (eds.) Proceedings of the Seventh Workshop on Building Educational Applications Using NLP. Association for Computational Linguistics. pp. 163–173. https://aclanthology.org/W12-2019/

von Garrel J, Mayer J (2023) Artificial intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. Humanit Soc Sci Commun 10:799. https://doi.org/10.1057/s41599-023-02304-7. Article

Wagner CS, Brahmakulam IT, Jackson BA, Wong A, Yoda, T (2001) Science & technology collaboration: Building capacity in developing countries? RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1357z0.html

Wang Y, Chen J (2013) Differences of English and Chinese as written languages and strategies in English writing teaching. Theory Pract Lang Stud 3(4):647–652. https://doi.org/10.4304/tpls.3.4.647-652

Warschauer M, Tseng W, Yim S, Webster T, Jacob S, Du Q, Tate T (2023) The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. J Second Lang Writ 62:101071. https://doi.org/10.1016/j.jslw.2023.101071

Xu X, Reed M (2021) The impact of internet access on research output: a cross-country study. Inf Econ Policy 56:100914. https://doi.org/10.1016/j.infoecopol.2021.100914

Zeng J, Yang J (2024) English language hegemony: retrospect and prospect. Humanit Soc Sci Commun 11(1):317. https://doi.org/10.1057/s41599-024-02821-z

Zhang S, Hasim Z (2023) Perceptions and coping strategies in English writing among Chinese study-abroad graduate students. Sage Open 13(3). https://doi.org/10.1177/21582440231184851

Author contributions

Arjun Prakash was responsible for the conceptualisation, methodology, data analysis, literature review, and the drafting and revision of the manuscript. Shruti Aggarwal contributed to manuscript review and editing, data curation, and final proofreading. Jeevan John Varghese contributed to the literature review and data curation. Joel John Varghese contributed technical support for data analysis. All authors have read and approved the final version of the manuscript. Arjun Prakash is the corresponding author and can be contacted at psarun.mangad@gmail.com.

Competing interests

The authors declare no competing interests.

Ethical approval

This study does not involve human participants, their data, or biological materials. All analyses are based on publicly available datasets that do not contain personally identifiable information. As such, ethical approval was not required according to institutional and international research guidelines.

Informed consent

The study does not involve human participants or the collection of personal or sensitive data. Therefore, informed consent was not applicable.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-05484-6.

Correspondence and requests for materials should be addressed to Arjun Prakash.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025