



ARTICLE



<https://doi.org/10.1057/s41599-024-03861-1>

OPEN

Elucidating university students' intentions to seek automated writing feedback from Grammarly: toward perceptual and systemic predictors

Yupeng Lin ¹ & Zhonggen Yu ²✉

Automated writing evaluation has gained popularity in technology-enhanced language education, while learners' technology acceptance of such tools has yet to be thoroughly investigated. Based on 487 university students' perceptions, this structural equation modeling study finds that traditional hypotheses in the Unified Theory of Acceptance and Use of Technology (UTAUT) can be extended to explain higher education students' Grammarly utilization. Students' performance and effort expectancy significantly predict their behavioral intentions to use Grammarly. Facilitating conditions and behavioral intentions significantly predict actual use behavior. Additionally, students' peer influence and trust in feedback significantly predict their performance expectancy. Peer influence, systemic interactivity, and personal investment significantly predict their effort expectancy. Students' willingness for e-learning and instructional support significantly predict facilitating conditions, while peer influence does not. The proposed model yields moderate explanatory power for Grammarly acceptance and use among higher education students ($R^2 = 19.6\% - 46.4\%$). Our findings can enlighten innovative technological designs for automated writing evaluation tools and digitalized foreign language writing instruction.

¹Faculty of Foreign Studies, Beijing Language and Culture University, Beijing, China. ²Department of English Studies, Faculty of Foreign Studies, Beijing Language and Culture University, Beijing, China. ✉email: 401373742@qq.com

Introduction

In the artificial intelligence era, natural language recognition, comprehension, and processing benefit numerous educational technologies, such as machine translation and automated writing evaluation (Huawei and Aryadoust 2023; Li 2023). Dominantly used for language education, automated writing evaluation tools are programs and platforms providing feedback and scoring mainly according to grammar and vocabulary based on textual materials (Liu et al. 2021). They bring multifarious benefits, improving students' learning experiences (Yao et al. 2021), enhancing teachers' and students' feedback literacy (Li 2023; Link et al. 2022), allowing prompt feedback, and reducing teachers' workload (Wang et al. 2022). Students get personalized recommendations from automated writing evaluation tools to enhance their language fluency, learning interests, and motivations (Guo et al. 2022). These tools complement traditional modes like peer assessment in foreign language education (Yao et al. 2021). Automated, teacher, and peer feedback show different focuses and influence language learning outcomes differently (Tian and Zhou 2020). Researchers are keenly interested in elucidating the effectiveness of automated writing evaluation tools and encouraging their applications in teaching practice. Such studies enrich the literature on assisting foreign language writing and teaching with automated writing evaluation tools (Li 2021).

Particular tools, such as *Grammarly*, *Pigai*, and *Criterion*, have contributed to related topics more efficiently than researching the general concept of automated writing evaluation (Ding and Zou 2024). Yang et al. (2023) set an example to investigate a famous artificial intelligence-empowered automated writing evaluation platform named *Pigai* in China and suggested students' improvements in the grammatical accuracy of English writing. In line with their methods, this study is motivated by the fact that *Grammarly* is a relatively less investigated automated writing evaluation tool and that research interest rises with the popularity of artificial intelligence-empowered learning tools (e.g., Chang et al. 2021; Fitria 2021). Founded in 2009, Grammarly has attracted over 70,000 business teams and 3000 educational institutions, collaborating with numerous stakeholders in educational and business sectors; it is also embracing artificial intelligence technologies and empowering its automated writing evaluation functions with pioneering techniques (Grammarly 2024; Ding and Zou 2024). Therefore, this platform has become increasingly influential among automated writing evaluation tools.

Researchers have probed into how Grammarly provides corrective feedback and students' perceptions (Koltovskaia 2020). However, few studies have thoroughly explored Grammarly as an automated writing evaluation tool in foreign language writing from the perspective of technology acceptance. Technology acceptance studies demonstrate their significance in revealing factors influencing educational technology use, which provides suggestions for teachers, educational administrators, and researchers in the digital era (e.g., Zhai and Ma 2022; Lin and Yu 2024a). Among various predictors of technology acceptance, perceptual and systemic indicators are popular in recent studies (Ismatullaev and Kim, 2024). For example, technological simplicity and systemic reliability significantly and positively predicted technology acceptance (Kikawa et al. 2022; Papakostas et al. 2023); perceived risks and support significantly impacted the intentions to use educational technologies (Al-Adwan et al. 2022; Lin and Yu 2023). However, various perceptual and systemic factors have yet to be specified and incorporated into acceptance and adoption research on automated writing evaluation tools. Hence, it is imperative to consider these aspects and elucidate technology acceptance and use of specific automated writing evaluation tools.

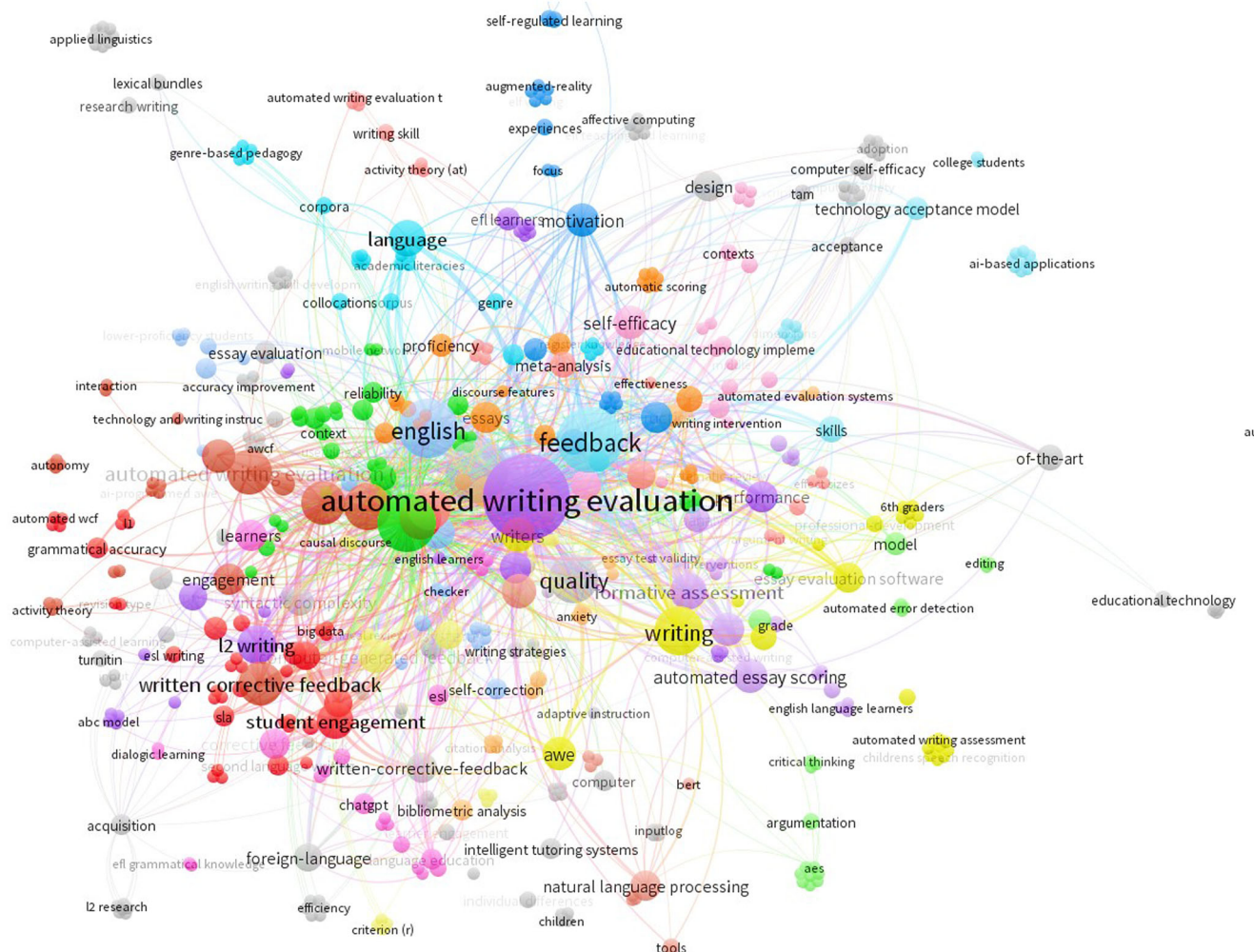
This study adopts structural equation modeling techniques to explore the acceptance and use of automated writing evaluation, with Grammarly as an example. We incorporate trust in feedback, peer influence, interactivity, personal investment, willingness for e-learning, and instructional support into the traditional Unified Theory of Acceptance and Use of Technology (UTAUT) model. We intend to establish how these innovative predictors introduced in this study can predict and explain technology acceptance in Grammarly utilization. The interrelationships among the variables can be tested through statistical approaches. We complement those relationships with learners' subjective evaluations of the advantages and disadvantages of Grammarly. In what follows, we first review the related literature on automated writing evaluation and UTAUT, followed by a proposed model containing new constructs and traditional variables (Section "Literature review"). Research methods will be described in Section "Methods". In the Section "Results", the path analysis will help test hypotheses among these variables, which will be informed by thematic analysis of the qualitative data. The extended Section "Discussion" will shed light on future language education assisted by automated writing evaluation tools and improve the functionalities and user interface designs of such platforms. Section "Conclusion" summarizes the major findings, limitations, and implications for future research and teaching based on our findings.

Literature review

Technology acceptance research based on automated writing evaluation tools. Technology acceptance research has been popular in various areas, including education, while the existing literature has yet to fully explore higher education students' adoption of automated writing evaluation tools despite their predominant advantages. On 19 April 2024, we conducted a literature search on the Web of Science Core Collection based on the following keywords in "Topic": "automated writing evaluation" OR "automated writing scoring" OR "automated writing feedback". The literature search led to 153 results initially analyzed through a keyword visualization and clustering analysis based on VOSviewer, a popular bibliometric study program created by van Eck and Waltman (2010). Figure 1 displays all 562 keyword items in the literature search outcomes. According to the occurrence frequencies, the most frequently used ten keyword items included automated writing evaluation (AWE) (occurrences = 100), feedback (occurrences = 44), students (occurrences = 38), English (occurrences = 35), accuracy (occurrences = 33), writing (occurrences = 22), quality (occurrences = 22), teacher (occurrences = 17), written corrective feedback (occurrences = 17), and impact (occurrences = 16).

The top ten keywords indicated that research on automated writing evaluation tools concentrated on its primary function of feedback. Such tools were designed to provide corrective feedback mostly on English writing and improve writing quality by identifying grammatical errors, adjusting sentence structures, and suggesting diverse synonyms (Zhai and Ma 2023; Gao 2021). It seemed that more studies took the students' perspectives in English learning contexts (e.g., Wei et al. 2023; Yang et al. 2023), while teacher involvement and perspectives (e.g., Link et al. 2022) were less investigated. Moreover, the top keywords suggested one of the most crucial criteria in automated writing assessment, i.e., "accuracy," which was applied to the quality of automated feedback and learner revisions (Kloppers 2023; Saricaoglu and Bilki 2021). Overall, examining the impacts of automated writing evaluation tools, researchers emphasized the quality assessment and assurance of foreign language writing (Ding and Zou 2024; Geng et al. 2024).

Based on the connections extended from the top keyword items, we identified four research trajectories of the existing



literature. Firstly, the tools were mostly incorporated into formative assessment to evaluate essay writing, which was represented by keywords like “formative assessment” and “formative feedback” (e.g., Foster 2019; Roscoe et al. 2017). Secondly, they could be investigated along with “syntactic complexity,” “genre,” and lexical features (e.g., Fan 2023; Lim et al. 2023). Thirdly, it was studied primarily in second and foreign language learning contexts, which was reflected in “EFL writing,” “EFL learner,” and “L2 writing” (e.g., Geng et al. 2024; Nguyen 2023). Fourthly, researchers grounded their studies on technology acceptance to explore many external factors associated with tool adoption, such as “computer self-efficacy,” “expectation,” “acceptance,” and “adoption” (Wei et al. 2023; Li 2021). However, studies on these tools based on technology acceptance were still scarce, and the factors influencing the adoption of automated writing evaluation tools needed more extensive investigations.

traditional Technology Acceptance Model. In another article, Li (2021) added confirmation and satisfaction to extend the psychological mechanisms behind accepting and adopting automated writing evaluation tools among students in China. Focusing on Grammarly use, Chang et al. (2021) confirmed that students with its assistance outperformed others with traditional intervention. Perceived motivation and enjoyment also positively influenced students' acceptance and adoption of automated writing evaluation tools (Nunes et al. 2022). Despite the popularity confirmed by numerous studies and reviews on such tools (Huawei and Aryadoust 2023; Wei et al. 2023; Huang et al. 2024), research on students' perceptions has yet to sufficiently clarify the influencing factors of acceptance and adoption.

Unified theory of acceptance and use of technology (UTAUT). The Unified Theory of Acceptance and Use of Technology was one of the prevalent technology acceptance models widely extended in educational research. This model was based on eight technology acceptance models, enhancing the explanatory power for technology acceptance in various contexts (Venkatesh et al. 2003). Lin and Yu's (2023) study extending the Technology Acceptance Model suggested that previous research on technology acceptance modeling was considered to combine two primary directions: (1) incorporating new constructs into established technology acceptance models and (2) extending technology acceptance models to new contexts, such as technologies, educational levels, and learning contents (e.g., Strzelecki 2024; Patil

and Undale 2023; Budhathoki et al. 2024). Extensive research enhanced the explanatory power of the technology acceptance models with new constructs (Al-Ghaith 2015; Lin and Yu 2023). Meanwhile, bringing specific external factors into technology acceptance models made contributions despite the cost of models' moderate explanatory power (e.g., Deng and Yu 2023; Wang et al. 2022; Lin and Yu 2024a). Researchers could delve into deeper and finer mechanisms of technology acceptance and innovatively explore these models from various areas.

Five critical variables in UTAUT were explored in this study to extend this model to higher education students' Grammarly utilization. Definitions of these traditional variables could be adapted from previous literature (e.g., Chang 2012): *Performance expectancy* (PE) refers to the degree to which higher education students believe that using Grammarly helps them with foreign language writing; *effort expectancy* (EE) refers to higher education students' perceived ease of using Grammarly to assist their foreign language writing; *facilitating conditions* (FC) refer to external conditions (such as institutional, technological, and relational ones) that allow higher education students to use Grammarly to assist their foreign language writing; *behavioral intentions* (BI) refer to higher education students' intentions and motivations to use Grammarly to assist their foreign language writing; *actual use behavior* (AUB) refers to higher education students' actual behavior of using Grammarly to assist their foreign language writing.

Traditional hypotheses in UTAUT were tested in various educational technologies. Proposed by Venkatesh et al. (2003), UTAUT contained six predicting variables with four additional moderators. According to their evaluation, the unified model outperformed its components, with the adjusted R^2 value reaching 69% (Venkatesh et al. 2003). Al-Adwan et al. (2022) incorporated three new constructs into the classical model to explore students' learning management system use, where the extended model yielded substantial explanatory power ($R^2 = 81.8\%$). Research on the traditional UTAUT model in new contexts could also clarify the acceptance and use of different educational technologies. Abbad's (2021) research on e-learning system acceptance was grounded on UTAUT without specifying the moderating effects, recording the explanatory power of 47.7% ~ 59.9%. The model was adopted to identify a disconfirmation effect in higher education students' acceptance of an early warning system, revealing the perceived challenges in applying such a technology (Raffaghelli et al. 2022). The extended UTAUT (UTAUT2) model was recently validated in explaining higher education students' ChatGPT utilization, with substantial explanatory power (Strzelecki 2024). In contrast, research extending UTAUT was still scarce in automated writing evaluation tools like Grammarly. We adapted the following four hypotheses from the traditional model to our research context:

H1: Higher education students' performance expectancy significantly predicts their behavioral intentions to use Grammarly.

H2: Higher education students' effort expectancy significantly predicts their behavioral intentions to use Grammarly.

H3: Higher education students' behavioral intentions to use Grammarly significantly predict their actual use behavior.

H4: Higher education students' facilitating conditions for using Grammarly significantly predict their actual use behavior.

External variables extending the Unified Theory of Acceptance and Use of Technology

Trust in feedback (TF). In previous studies on automated writing evaluation and feedback quality, *trust in feedback* refers to the degree to which learners believe automated feedback tools

provide valuable comments and suggestions (e.g., Ranalli 2021). Adapted to our research context, this concept measures how higher education students believe that suggestions provided by Grammarly are suitable and assistive for their foreign language writing tasks. Trust was derived from information system quality that largely accounted for users' perceived usefulness of and satisfaction with certain information technologies and systems (Kassim et al. 2012). A meta-analytical review disclosed previously established positive effects of trust on satisfaction, enjoyment, and continuance intentions (Mishra et al. 2023).

Trust in feedback was investigated in many contexts. Vorm and Combs investigated intelligent systems to establish the role of user trust in the Technology Acceptance Model (2022); Dautzenberg and Voß's study focused on trust in automation (2022). Previous researchers linked trust perceived in human-computer interaction with the perceived usefulness of technologies (e.g., Mou et al. 2017). This relationship revealed how the human-computer interaction indicator was associated with the expected benefits of using specific technologies. However, the effect has yet to be extensively validated in automated writing evaluation tools. Rationales existed that learners' trust in automated feedback could reflect how useful students considered Grammarly, which determined whether they were willing to adopt it in learning. To test this relationship in Grammarly use, this study proposed the following hypothesis:

H5: Higher education students' trust in feedback provided by Grammarly significantly predicts their performance expectancy.

Peer influence (PINF). *Social influence* refers to the degree to which people related to technology users think they should use certain technologies (Venkatesh and Morris 2000). Peer influence could be regarded as a sub-concept of social influence, which was seldom investigated in technology acceptance research (e.g., Yu et al. 2021). Yu and Yu (2019) specified social influence from its sources, included superior and peer influence to extend the Technology Acceptance Model, and suggested that superior and peer influence significantly predicted four traditional technology acceptance variables. Trivedi et al. (2022) found that peer influence significantly moderated the impacts of performance, perceived usefulness, and self on behavioral intentions to use social media among higher education students. Peer influence significantly moderated the impacts of perceived usefulness and ease of use on attitudes toward e-learning technology adoption in vocational education and training (Chatterjee et al. 2021).

In this study, the definition of peer influence was adapted from previous studies (Yu and Yu 2019), which considered the impact of higher education peers on students' Grammarly use. The role of peer influence could be supported by recent research on technology acceptance, but it was not investigated in the context of automated writing evaluation tools like higher education students' Grammarly utilization. As peer collaboration and communication demonstrated their increasing significance and competition became fierce in educational contexts (Horta and Li 2023), it was imperative to explore how peer influence played a role in technology acceptance. Therefore, we would like to test its relationships with the traditional variables in UTAUT by the following hypotheses:

H6: Peer influences on higher education students significantly predict their performance expectancy of using Grammarly.

H7: Peer influences on higher education students significantly predict their effort expectancy of using Grammarly.

H8: Peer influences on higher education students significantly predict their facilitating conditions of using Grammarly.

Interactivity (INT). *Perceived interactivity* highlights the extent to which users feel that they have the ability to influence the

functionalities and outcomes of information systems (Chang et al. 2018). Adapted from this definition, *interactivity* is defined in this study as the degree to which Grammarly efficiently responds to users' operations on the platform and users control Grammarly through proper operations to employ the expected functions. Interactive learning environments and tools gained solid theoretical foundations in experiential and simulation-based learning (Chang et al. 2023), while interactive features of educational technologies were still in development (Kolb 2014). Go et al. (2020) extended the traditional Technology Acceptance Model to an interactive one, suggesting a significant role of perceived interactivity in technology acceptance regarding consumption and commerce. More related to educational research, AL-Sayid and Kirkil (2023) delved into eight fine-grained human-technology interaction factors and established their non-linear relationships with the traditional Technology Acceptance Model. Interactive designs for educational technologies could facilitate learners' use and enhance their engagement in learning (e.g., Kennewell et al. 2008). Additionally, interactive learning modes enhanced students' performance and efficacy (Hwang and Chen, 2023).

Interactive information systems could provide timely feedback for learners in response to their operations. In this way, learners might find it easier to fulfill the expected functions by controlling the platforms. Additionally, the systemic, functional, and content features could be associated with technological and environmental conditions required for specific tasks like writing evaluations, influencing students' satisfaction with e-learning systems (Wu et al. 2010). Although the rationales existed for the impact of interactivity on perceived ease of using technologies, the relationship was underexplored through empirical studies on automated writing evaluation tools. With a keen interest in interactive designs for educational technologies, we intended to explore the role of interactivity through the following hypothesis:

H9: The interactivity of Grammarly significantly predicts higher education students' effort expectancy of using it.

Personal investment (PSI). *Personal investment* refers to the degree to which people are willing to put their resources (such as time and money) into particular tasks and items (Wang et al. 2022). This concept was adapted to our research context to examine how much students would devote to using Grammarly. This concept was derived from the Personal Investment Theory, which was proposed to explain students' activities in foreign language learning regarding motivation, goal-achieving, and expectation (King et al. 2019). It was also associated with commitment to academic goals in the Goal-setting Theory (Locke and Latham 2015). This concept was studied in some empirical research from the perspective of technology acceptance and educational technologies. Wang et al. (2022) adopted personal investment as an external variable to extend the Technology Acceptance Model, suggesting that it significantly contributed to students' perceived usefulness and behavioral intentions to use a mobile learning application. Taking the perspective of investment, researchers revealed that learners with different strategies in mobile-assisted language learning gained different returns on learning investment (Li and Wang 2024). However, inconsistent findings suggested that personal investment might not predict learners' continuance intentions in multimodal language learning (Huang et al. 2024).

Digital tools and online resources have become predominant intellectual property nowadays. Willingness to pay for online and digital learning styles demonstrated its importance (Berhan et al. 2023). It might determine what functions students could use with digital learning tools and how students' learning could benefit from paid digital learning tools and resources. However, this

concept was not thoroughly investigated in technology acceptance research of automated writing evaluation tools. With artificial intelligence integrated into automated writing evaluation tools, the paid functions of digital resources and tools could widen the differences in e-learning outcomes. Therefore, this study would investigate the following hypothesis:

H10: Higher education students' personal investment in Grammarly significantly predicts higher education students' effort expectancy of using it.

Willingness for e-learning (WEL). Previous investigations enriched the role of willingness in various contexts, such as willingness to communicate in digitalized learning environments or adopt innovative learning technologies (Zadorozhnyy and Lee 2023; Deng and Yu 2023). According to previous definitions, willingness reflected the attitudes and intentions toward specific tasks (e.g., Deng and Yu 2023). We intended to explore willingness for e-learning, as in Patil and Undale (2023). Educational technologies urged a transformation of learning to digital learning environments. However, students showed inconsistent attitudes toward digital learning tools and e-learning environments. Based on previous studies, willingness for e-learning was explored in our context, measuring higher education students' voluntariness to adopt a series of tools to form a digitalized workflow for learning.

Numerous studies observed students' proactive use of digital learning tools and its multifarious influencing factors, especially during the COVID-19 pandemic (Johnson et al. 2021; Yu and Yu 2023). In contrast, students might resist transforming their learning to new flows and media, even if these educational technologies and resources could greatly facilitate learning (Aboagye et al. 2021). Previous research also explored similar ideas in various concepts, for example, "personal innovativeness" in Deng and Yu's (2023) investigation of TikTok use in higher education, despite slight differences between their concept and willingness for e-learning in this study. Digital learning readiness might account for the conditions of e-learning, which was investigated in association with technology adoption intentions (Rafiee and Abbasian-Naghneh 2021). In the rapid transformation to a technology-enhanced education era, willingness for e-learning from the perspective of technology acceptance research could allow researchers to dive into learners' attitudes. The inconsistent findings in the existing literature and the scarce evidence in automated writing evaluation tool applications might confound further practice of learning and teaching. Therefore, we would test the following hypothesis:

H11: Higher education students' willingness for e-learning significantly predicts their facilitating conditions for using Grammarly.

Instructional support (IS). *Instructional support* measures students' accessible instruction and support regarding specific tasks (Mai et al. 2024; Gedrimiene et al. 2024). In our context, we explored perceived support through the instruction about how to use Grammarly to assist their foreign language writing successfully. This concept should be a component of facilitating conditions. Technological challenges dominantly prevented students from adopting digital learning tools, for example, in online learning contexts (Rasheed et al. 2020). Students and teachers might be bothered by similar difficulties in gaining or providing instruction on using educational technologies and unwillingness to shift to digital education practices (Bond et al. 2018). Students might seek resources for technological instruction, and the available instructional support could help them solve some problems when they used learning technologies. The perceived support mediated the relationship between the application of the

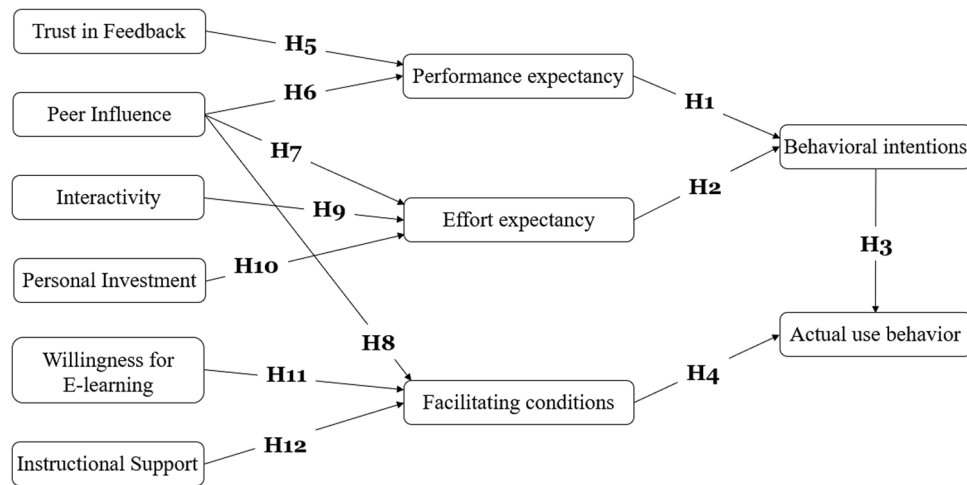


Fig. 2 A proposed model extending UTAUT to explain higher education students' acceptance and adoption of Grammarly. Note. Research hypotheses are numbered from H1 to H12 on the paths.

community of inquiry and students' learning outcomes (Sun and Yang, 2023).

Relevant instruction regarding students' use of educational technologies might significantly reduce the challenges and unwillingness to use digital learning tools (e.g., Schworm and Gruber 2012). When education turned from instructor-led to student-centered modes, self-scaffolding functions in e-learning tools could better satisfy personalized learning needs (Costley and Lange 2023). Despite its critical role, instructional support was still unclear in automated writing evaluation applications. We intended to test the positive impact of instructional support on facilitating conditions in Grammarly utilization through the following hypothesis:

H12: Instructional support significantly predicts higher education students' facilitating conditions of using Grammarly.

Research aims, questions, and a postulated model. The keyword visualization and systematic review above revealed four primary research paths in practicing automated writing evaluation tools. In line with technology acceptance research, this study intended to introduce some innovative factors that would explain the acceptance and adoption of such tools. The shortage of research exploring technology acceptance and adoption of Grammarly encouraged us to contribute to this topic by establishing the interrelationships underneath the proposed variables. Figure 2 demonstrates the interrelationships between six innovative external factors and five traditional concepts from UTAUT. Based on the above literature reviews, 12 research hypotheses were tested through this study. We aimed to explore how perceptual and systemic factors influenced higher education students' attitudes toward Grammarly use and their actual use behavior. We would answer two research questions (RQs) in this study:

RQ 1 (for quantitative phase): How do perceptual and systemic factors influence higher education students' acceptance and adoption of Grammarly?

RQ 2 (for qualitative phase): What are the primary advantages and disadvantages of Grammarly according to higher education students' perceptions?

Methods

Survey instrument design. We designed a comprehensive questionnaire survey to test the hypotheses, elicit subjective opinions, and explore the research questions. The questionnaire consisted of three sections. The first section collected informed consent,

age, gender, city, and educational stage. The second section measured variables, each with four or five statements. Eleven variables in this study were measured with 45 items. Specifically, the variables with the referred measurements were as follows: For the traditional variables within UTAUT, such as PE, EE, FC, BI, and AUB, we referred to previous validations and extensions of this model (e.g., Venkatesh et al. 2011; Abbad 2021; Lakhali and Khechine 2021; Hou and Yu 2023). For innovative variables that we incorporated into the traditional model, we adapted the measurements from other empirical studies, although they might not rely on UTAUT. Such variables and their references were as follows: IS (Lin and Yu 2023), TF, INT, PINF (Yu and Yu 2019), PSI (Wang et al. 2022), and WEL (Chang et al. 2017).

An additional attention test was set in the middle to examine whether participants were sufficiently attentive when responding to the questionnaire. The regular items measuring each variable were based on previous literature. In contrast, some variables were not explored in research on technology acceptance, where we demonstrated our innovative contributions to evaluating and incorporating them into our proposed model. The self-designed measurements based on literature could be established by content and statistical validation, demonstrating sufficient validity and reliability. Appendix A of this article provided a complete list of statements used to evaluate each variable. Participants rated each item with a five-point Likert scale, from one (strongly agree) to five (strongly disagree). The third section was an open-ended question, performing an indirect interview about the participants' perspectives on using Grammarly.

Research procedure. When the questionnaire was designed, two trained experts in technology-enhanced education and structural equation modeling research reviewed and approved the questionnaire. We adopted a popular online questionnaire survey platform in China (www.wjx.cn) to pose the questionnaire, allowing us to use a hyperlink or a QR code to distribute it easily. For data collection, rapid convenience sampling methods were used for this study. Convenience sampling methods refer to the data collection practice where researchers invite participants in an available population close to hand (Obilor 2023). Although they were not completely random in nature, the convenience sampling methods were widely adopted for large-sample structural equation modeling studies (e.g., Sun and Guo 2024). It was one of the most frequently adopted data collection methods in applied linguistic research (Dörnyei 2007).

This sampling method encouraged the researchers to spread the questionnaire through the Internet, including social media and instant messaging applications. The questionnaire survey could reach as many participants as possible in this way, and all possible participants who had related learning experiences and considered themselves eligible could contribute to this survey. Our data were primarily collected through WeChat groups, where we invited participants by sending information about this survey, and the participants could choose to participate and provide their data voluntarily without further intervention. The data sources also included other social networking communities and platforms popular in China, such as Tencent QQ, ResearchGate, and Xiaohongshu. The sampling approaches followed previous methods for educational research using structural equation modeling, such as Hou and Yu (2023) and Deng and Yu (2023). Data quality could also be tested with rigorous statistical approaches in the following sections.

We distributed the questionnaire before the sample size reached a good level for further analysis. Previous studies widely suggested a sample of over 300 as a good one and over 500 as a very good one for structural equation modeling (Raza et al. 2020). Other researchers practiced the “ten times” principle, suggesting that the sample size should be more than ten times the measurements, i.e., 450 in this study (Hair et al. 2014; Hidayatur-Rehman and Ibrahim, 2023). We collected 548 responses from 11 April to 3 November 2023. According to informed consent (two participants disagreed), attention test results (53 participants failed), and educational levels (six participants were identified as teachers or workers), we obtained 487 valid records for the statistical analysis. The valid responses formed an adequate sample size for our statistical analysis.

After data collection, the survey platform allowed us to export the data as a CSV file, which was used first to reveal the demographic information about our participants and to test the validity and reliability of our measurement. Assuring that the measurements were valid and reliable, we used SmartPLS 4.1.0.3 to test our proposed model and hypotheses with the collected data (Ringle et al. 2024). SmartPLS is a program based on partial least squares structural equation modeling (PLS-SEM). We could use the “PLS-SEM algorithm” to calculate the construct validity and reliability, discriminant validity, collinearity statistics measured by the variance inflation factor (VIF), and R^2 values; the “Bootstrapping” function could calculate path coefficients and their statistical significance; the “PLSpredict” function could evaluate the model’s predictive power measured by Q^2 values. Reporting of these indices referenced the existing publications like Gao et al. (2024), and those statistics would be explained in the following sections. The qualitative results from the open-ended question were analyzed through thematic classification and analysis regarding the advantages and disadvantages of Grammarly by both authors.

Statistical approaches

Eligible participant selection and normality test. After data collection, we first excluded participants who did not provide informed consent to the researchers since the survey questionnaire would end if the participants disagreed to participate. Focusing on higher education students’ acceptance and use of Grammarly, we also excluded participants identifying as “Teachers and workers”. According to the attention test, participants who did not choose “Agree” as required were considered not attentive when responding to the questionnaire, and their responses were excluded for more accurate estimates. These eligible responses should be tested for the normality of distribution, as Kline (2015) suggested that multivariate normality should be

assumed for structural equation modeling with maximum likelihood estimation. Kline (2015) provided the guideline for the normality test: When the absolute value of skewness was smaller than 3.000 and that of kurtosis smaller than 10.000, the normality of the dataset should be acceptable; otherwise, the distribution would be extremely skewed. The normality of each item was evaluated with SPSS 26.0.

Validity and reliability tests. We tested the validity and reliability of the questionnaire. First, we calculated factor loadings using items for each variable with SPSS. Factor loadings were calculated with the principal component method when the fixed number of factors was set as 11. We adopted the varimax method for factor rotation. A factor loading larger than 0.500 could show that the item measured the corresponding factor satisfactorily (Pham et al. 2019). In contrast, items with factor loadings smaller than 0.500 would indicate unsatisfactory validity and reliability due to poor designs. We removed items with low factor loadings from the measurement model for subsequent statistical analyses to get better validity and reliability. Internal consistency (Cronbach’s α) could be considered excellent when $\alpha > 0.900$, good when $0.800 \leq \alpha < 0.900$, or acceptable when $0.700 \leq \alpha < 0.800$ (Hutcheson and Sofroniou 1999).

We adopted a Kaiser-Meyer-Olkin (KMO) test and Bartlett’s test of sphericity. When the KMO value was greater than 0.750 and the p -value in Bartlett’s test was significant, the dataset would be suitable for further factor analysis and structural equation modeling (e.g., Zhang and Liu 2022). The average variance extracted (AVE) and composite reliability (CR) were also calculated. According to Fornell and Larcker’s criteria (1981), $AVE > 0.500$ and $CR > 0.500$ could demonstrate good validity and reliability of the questionnaire. The Fornell-Larcker method would be adopted to establish the discriminant validity based on AVE and CR values, testing whether the variables were statistically distinguishable (Fornell and Larcker 1981). Another approach was the Heterotrait-Monotrait ratio (HTMT) of correlations proposed by Henseler et al. (2014). For this method, a threshold was proposed by Kline (2015), suggesting that values smaller than 0.850 would demonstrate good discriminant validity.

Common method bias was frequently mentioned in various measurements and originated from the measurement designs in structural equation modeling studies (Kock 2015). Not to mention the theoretical and mathematical significance of testing such bias, a successful approach to identifying it in structural equation modeling research was to observe the variance inflation factor (VIF) values (Kock 2015). In other studies, researchers exempted their measurement models from the collinearity issue with this indicator, for which Hair et al. (2014) set a criterion of $VIF < 5.000$, or more ideally, $VIF < 3.000$ (Hair et al. 2019). According to Kock (2015), the low VIF values from a full collinearity test should indicate that the measurement model was not contaminated by common method bias. In addition, the correlation matrix of the measurement model could also be used, where correlation coefficients lower than 0.900 demonstrated the absence of common method bias (Lowry and Gaskin 2014).

Explanatory power of the model and path analysis. The assured validity and reliability of the questionnaire would demonstrate good measurements of the proposed variables. We used the collected data to evaluate to what extent the exogenous variables could explain and predict the endogenous ones. A common statistic used for explanatory power is the R^2 value. Ozili (2023) suggested three critical ranges reflecting different explanatory power in social sciences, i.e., $0 < R^2 < 0.100$ (too low), $0.100 \leq R^2 \leq 0.500$ (acceptable when most explanatory variables were statistically significant), and $R^2 > 0.500$ (significant). Sarstedt

Table 1 Participants' demographic information of valid records for statistical analysis.

Item	Type	Frequency	Percentage
Gender	Female	356	73.101%
	Male	131	26.899%
Educational level	Undergraduate	136	27.926%
	Master's student	328	67.351%
	Doctoral student	23	4.723%
Age	18–22 years old	150	30.801%
	22–25 years old	267	54.825%
	25–30 years old	62	12.731%
	Over 30 years old	8	1.643%
Total participants		487	

and Mooi (2014, p.211) adopted a more general principle: “ R^2 values of 0.75, 0.50, or 0.25” should be considered substantial, moderate, or weak.

Q^2 was a model predictive power indicator, which could be calculated by the PLSpredict function in SmartPLS. Q^2 values greater than zero allowed the researchers to compare the mean absolute error (MAE) or the root mean square error (RMSE) with the benchmarks calculated by the linear regression model (LM) (Hair et al. 2021). In PLS-SEM analysis, high predictive power requires all indicators to have lower RMSE (or MAE) values than the naive LM benchmark; medium predictive power is achieved when most indicators have smaller prediction errors than the LM; low predictive power is indicated if only a minority of indicators show lower PLS-SEM prediction errors than the LM; the lack of predictive power is evident if none of the indicators in PLS-SEM have lower prediction errors than the LM (Hair et al. 2021).

Path analysis would examine the effect of one variable on another that could be measured by a path coefficient (β) and an f^2 value at the significant level of $p < 0.050$. The path would demonstrate a positive effect when $\beta > 0$ and a negative one when $\beta < 0$. Effect sizes would be determined by different ranges of the absolute values of β . When the absolute values of β fell into 0 ~ 0.100, 0.100 ~ 0.300, 0.300 ~ 0.500, or over 0.500, the effect size of the path would be weak, modest, moderate, or strong (Zhang and Liu 2022). According to Cohen (1988), f^2 values would indicate large ($f^2 > 0.350$), medium ($f^2 > 0.150$), or low effect sizes ($f^2 < 0.020$).

Results

Demographic information and normality test of each item.

Table 1 displays the demographic information of 487 eligible responses that could be considered a good sample for structural equation modeling. The subsequent statistical analysis would be based on these valid records. Due to our personal network and research areas, participants were dominantly females, which was common in arts, humanities, and social sciences majors in China. To reflect higher education students' attitudes toward Grammarly utilization more comprehensively, the survey also invited students from natural science majors during our data collection, such as engineering, mathematics, and medical science. Our survey questionnaire was actively responded to by master's students and undergraduates, with only a few doctoral students. Participants were from 28 provinces or regions of China and identified as regular Grammarly users, among which the largest proportion came from Beijing (19.10%). The low proportion of each province and region indicated the diverse sources and the robust representativeness of our participants. Responses to each item were tested for the normality of distribution. The skewness of all items was [−1.130, −0.128], and the kurtosis was [−0.684, 3.342]. The absolute values of skewness and kurtosis were smaller than 3.000

Table 2 Measurement model assessment.

Variable	Item	Factor loading	VIF	Cronbach's α	CR	AVE
AUB	AUB1	0.751	2.686	0.919	0.922	0.805
	AUB2	0.759	2.644			
	AUB3	0.768	3.801			
	AUB4	0.727	3.241			
BI	BI1	0.646	2.647	0.899	0.899	0.767
	BI2	0.566	2.136			
	BI3	0.576	2.471			
	BI4	0.646	3.071			
EE	EE1	0.628	2.006	0.893	0.895	0.700
	EE2	0.570	1.878			
	EE3	0.791	3.161			
	EE4	0.671	2.559			
	EE5	0.761	2.306			
FC	FC1	0.720	2.205	0.858	0.864	0.700
	FC2	0.814	1.838			
	FC3	0.781	2.032			
	FC4	0.613	1.891			
INT	INT1	0.686	2.122	0.846	0.846	0.764
	INT2	0.714	2.236			
	INT3	0.725	1.849			
IS	IS1	0.611	2.168	0.881	0.887	0.737
	IS2	0.769	2.680			
	IS3	0.618	2.593			
	IS4	0.569	1.895			
PE	PE1	0.726	1.968	0.829	0.830	0.745
	PE2	0.768	1.935			
	PE3	0.646	1.803			
PINF	PINF1	0.778	2.646	0.884	0.890	0.743
	PINF2	0.749	2.474			
	PINF3	0.664	2.049			
	PINF4	0.651	2.394			
PSI	PSI1	0.756	2.275	0.869	0.869	0.718
	PSI2	0.790	2.533			
	PSI3	0.712	2.087			
	PSI4	0.592	1.766			
TF	TF1	0.610	2.161	0.882	0.885	0.738
	TF2	0.722	2.259			
	TF3	0.702	2.436			
	TF4	0.707	2.382			
WEL	WEL1	0.778	3.052	0.912	0.920	0.791
	WEL2	0.820	2.820			
	WEL3	0.808	3.119			
	WEL4	0.799	2.491			

AUB actual use behavior, BI behavioral intentions, EE effort expectancy, FC facilitating conditions, INT interactivity, IS instructional support, PE performance expectancy, PINF peer influence, PSI personal investment, TF trust in feedback, and WEL willingness for e-learning. Statistic indices in this table included CR composite reliability, AVE average variance extracted, and VIF variance inflation factor.

and 10.000, respectively, indicating that the dataset was normally distributed. The normality test results satisfied the preconditions for factor analysis.

Validity and reliability of the survey questionnaire. The KMO value was 0.960 > 0.750, and Bartlett's test was statistically significant ($p < 0.001$), which satisfied the preconditions of factor analysis. According to the confirmatory factor analysis results, we deleted INT4 and PE4 due to their low factor loadings (<0.500), after which the rest items gained satisfactory factor loadings to 11 different variables, and the preconditions based on the KMO value and Bartlett's test remained supported. Table 2 demonstrates that the factor loadings were all greater than 0.500. Examining variables with their items, we found satisfactory internal consistency with Cronbach's $\alpha > 0.500$, composite

Table 3 Discriminant validity test results using the Heterotrait-Monotrait ratio (HTMT) of correlations.

	AUB	EE	FC	INT	IS	BI	PE	PINF	PSI	TF	WEL
AUB											
EE	0.616***										
FC	0.550***	0.773***									
INT	0.558***	0.586***	0.441***								
IS	0.724***	0.600***	0.471***	0.692***							
BI	0.710***	0.656***	0.536***	0.671***	0.670***						
PE	0.502***	0.775***	0.610***	0.499***	0.575***	0.665***					
PINF	0.668***	0.568***	0.407***	0.611***	0.801***	0.654***	0.514***				
PSI	0.657***	0.592***	0.388***	0.659***	0.722***	0.566***	0.490***	0.744***			
TF	0.606***	0.574***	0.425***	0.798***	0.713***	0.736***	0.569***	0.615***	0.655***		
WEL	0.482***	0.482***	0.364***	0.558***	0.556***	0.766***	0.551***	0.548***	0.451***	0.641***	

AUB actual use behavior, BI behavioral intentions, EE effort expectancy, FC facilitating conditions, INT interactivity, IS instructional support, PE performance expectancy, PINF peer influence, PSI personal investment, TF trust in feedback, WEL willingness for e-learning.
*** $p < 0.001$.

Table 4 Discriminant validity test results using the Fornell-Larcker method.

	AUB	EE	FC	INT	IS	BI	PE	PINF	PSI	TF	WEL
AUB	0.897										
EE	0.560***	0.837									
FC	0.491***	0.681***	0.837								
INT	0.493***	0.510***	0.376***	0.874							
IS	0.653***	0.537***	0.418***	0.601***	0.858						
BI	0.648***	0.591***	0.473***	0.586***	0.599***	0.876					
PE	0.439***	0.669***	0.519***	0.418***	0.494***	0.574***	0.863				
PINF	0.606***	0.509***	0.360***	0.529***	0.707***	0.586***	0.442***	0.862			
PSI	0.587***	0.524***	0.338***	0.566***	0.634***	0.503***	0.417***	0.652***	0.847		
TF	0.547***	0.512***	0.371***	0.688***	0.633***	0.658***	0.489***	0.546***	0.574***	0.859	
WEL	0.446***	0.441***	0.326***	0.491***	0.505***	0.695***	0.481***	0.495***	0.406***	0.578***	0.889

AUB actual use behavior, BI behavioral intentions, EE effort expectancy, FC facilitating conditions, INT interactivity, IS instructional support, PE performance expectancy, PINF peer influence, PSI personal investment, TF trust in feedback, WEL willingness for e-learning.
*** $p < 0.001$.

reliability (CR) > 0.500, and average variance extracted (AVE) values > 0.500.

The dataset was exempted from common method bias through the collinearity test. Collinearity test results showed that VIF < 3.000 was supported for most items, with several acceptably close to 3.000 and much smaller than 5.000. Meanwhile, the results could also be interpreted according to Kock's (2015) method of identifying common method bias in partial least squares structural equation modeling techniques. VIF values lower than 5.000 suggested that the measurement model in this study was free from obvious common method bias. Discriminant validity was tested using two methods in this study. First, the Heterotrait-Monotrait ratio of correlations were all smaller than 0.850. Table 3 demonstrates that the discriminant validity of this survey reached an acceptable level. The results supported that variables measured with the survey questionnaire could be distinguished statistically.

Then, the Fornell-Larcker method was also adopted to test the discriminant validity. Table 4 places the square roots of average variance extracted values in bold on the diagonal of this matrix, while inter-construct correlation coefficients (r) were displayed below between different variables. The square roots were larger than the correlation coefficients on their corresponding rows and columns. As such, we could support the discriminant validity in this study through the Fornell-Larcker method. This matrix also displays the correlations between variables, and significant correlations ($r \geq 0.500$) could be identified in 34 pairs in the following table. The three largest correlation coefficients could be found between (1) peer influence and instructional support ($r = 0.707$), (2) willingness for e-learning and behavioral intention ($r = 0.695$), and (3) systemic interactivity and trust in feedback ($r = 0.688$). All correlation coefficients in Table 4 were smaller than 0.900, which corroborated that the measurement model in our study was free from common method bias.

Path analysis and research hypothesis testing (RQ1). The above sections demonstrated that validity and reliability were acceptable for this study. Table 5 illustrates the full results of structural model estimates and the model's explanatory and predictive power. According to path analysis, we accepted all research hypotheses except H8: Peer influence could not significantly predict facilitating conditions. According to the f^2 values of the supported paths, a large effect size was identified in H3 ($f^2 = 0.415$) and a small effect size in H11. The other paths demonstrated moderate effect sizes. Measured by R^2 values, the proposed model could explain 19.6%–46.4% of variances in traditional variables in UTAUT regarding higher education students' use of Grammarly to assist their foreign language writing. The explanation power for the latent variables was acceptably moderate. Table 6 illustrates that measured by Q^2 , RMAS, and MAE values, the model had nearly medium predictive power for the outcome variables (i.e., PE, EE, FC, BI, and AUB), according to Hair Jr et al.'s criteria (2021). In short, the model had acceptably good explanatory and predictive power for higher education students' Grammarly utilization.

Thematic analysis of qualitative results (RQ2)

Automated suggestions about grammar and vocabulary. The advantages and disadvantages filtered in the qualitative phase primarily contained three aspects. Most dominantly, participants perceived benefits in the reduced workload for correcting grammatical mistakes. They affirmed that Grammarly greatly assisted them by identifying and correcting grammatical errors that they were unaware of. For example, "Non-native speakers may be insensitive to particular types of mistakes and grammatical rules. Grammarly addresses those issues effectively" (Participant 66). They found Grammarly an efficient writing tool to unify punctuation uses, which saved much time and effort, especially for academic purposes. "It helps me better comply with writing standards for academic purposes" (Participant 342).

Table 5 Results of structural model estimates.									
H.	Path relationship	Std. Beta	Std. Err.	t-value	p-value	95% CI	VIF	f ²	Q ² predict
Direct effects:									
H1	PE → BI	0.324	0.071	4.552	***	[0.184, 0.459]	1.810	0.098	0.392
H2	EE → BI	0.374	0.063	5.962	***	[0.251, 0.496]	1.810	0.130	
H3	BI → AUB	0.535	0.049	10.881	***	[0.436, 0.627]	1.288	0.415	0.329
H4	FC → AUB	0.238	0.049	4.833	***	[0.146, 0.336]	1.288	0.082	
H5	TF → PE	0.353	0.053	6.680	***	[0.245, 0.455]	1.425	0.122	0.274
H6	PINF → PE	0.249	0.043	5.820	***	[0.166, 0.334]	1.425	0.061	
H7	PINF → EE	0.218	0.049	4.418	***	[0.122, 0.317]	1.863	0.040	
H8	PINF → FC	0.091	0.063	1.441	0.150	[−0.030, 0.217]	2.107	0.005	0.356
H9	INT → EE	0.262	0.047	5.600	***	[0.169, 0.352]	1.574	0.069	0.182
H10	PSI → EE	0.234	0.050	4.658	***	[0.140, 0.337]	1.974	0.044	
H11	WEL → FC	0.138	0.055	2.505	0.012*	[0.029, 0.247]	1.414	0.017	
H12	IS → FC	0.284	0.058	4.923	***	[0.174, 0.398]	2.134	0.047	
Indirect effects:									
	EE → BI → AUB	0.200	0.038	5.247	***	[0.128, 0.277]			
	IS → FC → AUB	0.068	0.024	2.872	0.004**	[0.031, 0.122]			
	PINF → PE → BI	0.081	0.025	3.214	0.001**	[0.038, 0.134]			
	INT → EE → BI	0.098	0.026	3.785	***	[0.052, 0.152]			
	PINF → FC → AUB	0.022	0.016	1.345	0.179	[−0.007, 0.056]			
	TF → PE → BI	0.115	0.036	3.213	0.001**	[0.052, 0.189]			
	PE → BI → AUB	0.174	0.042	4.120	***	[0.092, 0.259]			
	PINF → EE → BI	0.081	0.023	3.509	***	[0.041, 0.131]			
	PSI → EE → BI	0.087	0.024	3.606	***	[0.045, 0.142]			
	WEL → FC → AUB	0.033	0.014	2.266	0.023**	[0.007, 0.064]			
	PINF → EE → BI → AUB	0.044	0.013	3.362	0.001**	[0.021, 0.072]			
	PSI → EE → BI → AUB	0.047	0.014	3.403	0.001**	[0.023, 0.077]			
	PINF → PE → BI → AUB	0.043	0.014	3.006	0.003**	[0.020, 0.075]			
	TF → PE → BI → AUB	0.061	0.020	3.050	0.002**	[0.027, 0.105]			
	INT → EE → BI → AUB	0.052	0.015	3.562	***	[0.026, 0.084]			
AUB actual use behavior, BI behavioral intentions, EE effort expectancy, FC facilitating conditions, INT interactivity, IS instructional support, PE performance expectancy, PINF peer influence, PSI personal investment, TF trust in feedback, WEL willingness for e-learning. H _i hypothesis, CI confidential interval. Std. Beta standardized beta (path coefficient), Std. Err. standard error value, i.e., STDEV in PLS-SEM bootstrapping results. *p < 0.05, **p < 0.01, and ***p < 0.001.									

Table 6 Predictive power estimates with PLSpredict.

Variable	Item	Item Q ² predict	PLS-SEM		LM	
			RMSE	MAE	RMSE	MAE
AUB	AUB1	0.253	0.844	0.675	0.792	0.567
	AUB2	0.243	0.970	0.800	0.893	0.679
	AUB3	0.282	0.872	0.714	0.775	0.557
	AUB4	0.284	0.918	0.759	0.827	0.598
BI	BI1	0.273	0.643	0.447	0.559	0.398
	BI2	0.313	0.678	0.505	0.596	0.437
	BI3	0.301	0.651	0.473	0.581	0.424
	BI4	0.312	0.651	0.470	0.537	0.387
EE	EE1	0.206	0.683	0.515	0.664	0.496
	EE2	0.315	0.706	0.562	0.722	0.567
	EE3	0.232	0.745	0.592	0.744	0.581
	EE4	0.278	0.706	0.549	0.699	0.535
FC	EE5	0.200	0.745	0.578	0.756	0.590
	FC1	0.159	0.758	0.564	0.764	0.580
	FC2	0.089	0.779	0.587	0.785	0.595
	FC3	0.134	0.838	0.662	0.854	0.682
PE	FC4	0.125	0.775	0.588	0.782	0.604
	PE1	0.218	0.654	0.486	0.655	0.486
	PE2	0.167	0.717	0.538	0.708	0.533
	PE3	0.223	0.692	0.516	0.707	0.521

However, they also highlighted that Grammarly sometimes provided inaccurate corrections for some errors, such as articles and verb-preposition collocations. Other participants suggested that the advice for academic writing might need to be more content-specific. They argued, “*The suggestions for different genres and purposes should be improved to distinguish my writing for formal, informal, academic, and general purposes. Although the options are provided, I don’t see many differences*” (Participant 235). Furthermore, the participants noted that Grammarly enhanced the diversity of vocabulary use and improved the accuracy of spelling and collocation. “*I can use it to enrich my expressions. However, the synonyms provided might be rigid and inappropriate in the context. The suggestions need my additional efforts to determine which synonyms fit my sentences.*” (Participant 217). Another participant (105) suggested, “*I need a dictionary and more exemplary sentences integrated into the platform to clarify the suggested usages.*”

User interface. They found the user interface to be difficult to manipulate and unstable. For some participants, the systems “*refreshed multiple times and might break down when there were too many errors in the texts*” (Participant 173). The configuration of the website “*needed to be simplified*” (Participants 69 and 137). Participants also indicated their need to use this platform on multiple devices, while they thought the plug-ins provided for Word Office and the websites were unstable and lowered their working effectiveness due to significant delays and slow responses. They considered the mobile version a considerate extension, but the user interface was inconvenient to use, with occasionally unstable and slow responses. Additionally, the participants considered the cost of the premium version so high that it prevented their use.

Incorporation of artificial intelligence. Participants shared the belief that artificial intelligence chatbots like ChatGPT now had the potential to outperform Grammarly. They pointed out that Grammarly “*lacked quick updates to its language model, failed to learn and adapt to users’ language use patterns, and struggled to recognize emerging popular words*” (Participant 24). They also noticed that the suggested synonyms provided by Grammarly

were not always appropriate for the given context. In contrast, large language models with generative functions allowed users to personalize their preferred revision suggestions. Artificial intelligence technologies could also provide suggestions on logic flow and coherence, while “*Grammarly seemed not to provide such suggestions on the level of the entire text*” (Participant 295). The participants appealed to further integration of artificial intelligence technologies into Grammarly.

Discussion

Extending traditional hypotheses in UTAUT to Grammarly. Paths adapted from the traditional UTAUT model (H1 to H4) can be supported in explaining higher education students’ Grammarly utilization. Thus, the findings in this study provide solid evidence for extending UTAUT to a new context, where a specific automated writing evaluation tool is used to assist foreign language writing. Considering similar variables conceptualized in the traditional Technology Acceptance Model, the findings in this study are consistent with Zhai and Ma’s (2022) modeling results of automated writing evaluation among Chinese college students, although they focused more on how systemic characteristics influenced students’ acceptance and use of such tools from technological sophistication, a rather compound and complex concept.

We also find consistent results about technology acceptance and use with previous validation of the traditional UTAUT model in various educational technologies and contexts, for example, learning management systems (Al-Adwan et al. 2022). Performance expectancy demonstrates higher education students’ expectation of benefits and system usefulness when Grammarly assists foreign language writing; effort expectancy corresponds to their perceived ease of using this platform or program. These two dimensions capture the primary motivators for students to use Grammarly. Thus, the relationship between learning motivation and automated feedback tools in previous literature can be extended to the sources of motivation from students’ perspectives (e.g., Yao et al. 2021). Based on the close relationships among technology acceptance models, the validated traditional hypotheses of UTAUT can also support theories such as UTAUT-3, Technology Continuance Theory (TCT), and Value-based Adoption Model (VAM) (Farooq et al. 2017; Pasupuleti and Thiyyagura 2024; Kim et al. 2007).

The facilitating conditions examined in this study significantly predict students’ actual use behavior of Grammarly (H4), presumably because they allow students to avoid many challenges and difficulties when using Grammarly. This explanation can be supported by previously established challenges on students’ side in technological and learning aspects when automated writing evaluation tools are incorporated into practice (Deeva et al. 2021). Technical challenges have inspired researchers to explore technology readiness, for example, under the Theory of Planned Behavior (Sungur-Gül and Ateş 2021). The mediating effect of behavioral expectations has been found significant in the effects of facilitating conditions on actual use behavior (Venkatesh et al. 2008), supporting the above reasoning for H4.

Interestingly, H1, H2, and H4 demonstrate smaller path coefficients (i.e., effect sizes) than H3. This reflects the differences between behavioral intentions and the actual use of Grammarly. A possible reason is that students still have concerns regarding incorporating Grammarly into their learning practice, even when they understand the functionality and usefulness of this tool. This is a common dilemma found in other educational technologies in various application contexts, for example, in a study of automated writing evaluation tools used in primary education (Wang et al. 2020). Based on our qualitative data in this survey, challenges

may include the high cost for full functions on this platform, some unsatisfactory suggestions on revision, and students' actual needs when foreign language writing tasks are not frequent in their learning presently. Additionally, the necessary skills and knowledge for digital learning are critical to the successful digitalization of higher education. To better prepare students for advancements in educational technologies, the Technological, Pedagogical, and Content Knowledge (TPACK) framework demonstrates its significance among teachers (Abubakir and Alshaboul 2023).

New constructs explaining Grammarly utilization. Most external factors explored in this study significantly predict higher education students' acceptance and use of Grammarly based on UTAUT. Trust in feedback quality perceived by higher education students can be a dominant aspect of their performance expectancy in using Grammarly (H5), determining the benefits students can gain from it. In a traditional peer assessment context, Rotsaert et al. (2018) established the critical role of feedback quality in encouraging secondary education students to accept peer feedback. Considering automated feedback provided by Grammarly, trust in feedback reflects students' subjective evaluation and perception of this automated writing evaluation tool. The result is consistent with Zhai and Ma's (2022) exploration of contributing factors to students' perceived usefulness. Many factors have been associated with feedback quality and students' trust in automated evaluation and feedback regarding foreign language writing, such as accuracy and reliability (Zhai and Ma 2022). Some negative findings can also be the evidence, indicating that surface-level feedback by automated evaluation reduced students' engagement and their incorporation of such tools into learning practice (Tian and Zhou 2020).

Peer students' influence contributes to performance and effort expectancy (H6 and H7), probably through recommendation and exemplification. Students may feel less challenged or at risk if their peers successfully use and recommend such learning tools. The explanation is grounded on the peer learning theories, where researchers established knowledge construction and skill acquisition as a process within socio-cultural interaction (Meschitti 2019). The results are consistent with Yu and Yu's findings (2019) in mobile-assisted learning. However, peer influence may have inconsistent impacts on students' acceptance and use of Grammarly. Based on H8, peer influence cannot significantly predict their facilitating conditions of using Grammarly. One possible reason is that excessive peer influence brings peer pressure that discourages students from easy and comfortable use of this tool. Alternative feedback of higher quality may also prevent students from using automated writing evaluation tools (Tian and Zhou 2020). This finding is consistent with Herting et al. (2023), who found an insignificant predicting effect of social influences on higher education students' intentions to use PowerPoint. Comments on Grammarly are inconsistent in our study and depend on subjective evaluations and specific educational needs, presumably leading to the insignificant effect of peer influence.

The systemic interactivity perceived in higher education students' use of Grammarly and personal investment significantly predict their effort expectancy (H9 and H10). Zhai and Ma (2022) explored the technological characteristics of automated writing evaluation, suggesting that the technological sophistication of the system was significantly related to students' perceived ease of use. Their systemic characteristics encompassed interactivity, flexibility, and a sense of control. In contrast, this study concentrates on the interactive designs of Grammarly and finds results congruent with their established contribution from a more

general perspective. Personal investment is a significantly positive predictor of effort expectancy. As this construct includes the investment of money and time, the paid premium functions may allow students to gain better user experiences, and more time invested promotes their proficiency in understanding and using automated writing feedback, hence stronger perceived usefulness and ease of use. From the theoretical perspective, this study specifies that among limited personal resources, the investment of time and money can result in easy utilization and strong use intention, which points out the challenges and opportunities of wider applications. This established effect (H10) may become more stable if learners' concerns about the high cost of the premium version can be addressed, or cheaper substitutes with better functionality will result in a stronger impact of effort expectancy in alternative tools.

Willingness for e-learning and instructional support significantly predict higher education students' facilitating conditions of using Grammarly (H11 and H12). Students' academic experiences and expectations contributed to their understanding of academic work, further encouraging them to exploit digital learning tools, as Lin and Yu (2023) established. The Expectation-Confirmation Model has been validated across educational technologies and learning contexts, suggesting that satisfying the expectations of using certain educational technologies leads to stronger perceived usefulness and the intentions to use them (Obeid et al. 2024; Bhattacharjee 2001). Extending confirmation, e-learning experiences, and academic expectations, students' willingness for e-learning reflects how they perceive the benefits of Grammarly in assisting their writing based on their experiences and expectations (Fu et al. 2022). Instructional support helps students avoid technical challenges and ensure that they can use the platform successfully. The findings of this study establish that those two aspects dominantly explain students' facilitating conditions for using Grammarly.

Conclusion

Major findings. This structural equation modeling research explored external factors from perceptual and systemic features influencing higher education students' acceptance and use of Grammarly to assist foreign language writing based on the traditional Unified Theory of Acceptance and Use of Technology. We found that traditional hypotheses were successfully extended to explain the acceptance and use of Grammarly. Students' performance and effort expectancy significantly predicted their behavioral intentions to use Grammarly. Facilitating conditions and behavioral intentions significantly predicted the actual use behavior of Grammarly. Additionally, external factors extended the traditional model. Students' trust in feedback and peer influence significantly predicted their performance expectancy. Peer influence, perceived interactivity, and personal investment significantly predicted their effort expectancy. Willingness for e-learning and instructional support significantly predicted the facilitating conditions of using Grammarly, while peer influence did not. Participants perceived dominant advantages in prompt and accurate feedback, while the inaccuracy in particular circumstances and the high cost resulted in significant concerns. The participants called for further incorporating artificial intelligence into the platform to enhance the accuracy and quality of automated feedback.

Limitations. We must acknowledge some limitations of this study. First, the outcome variables in the proposed model only yielded about 20%–50% explanatory power, suggesting that other variables not included in the research model may also have

significant explanatory and predictive power. However, it was inevitable that some innovative and specific variables explored in technology acceptance models came at the cost of models' low explanatory power. In social science research, small explanatory power (R^2 of 10%–50%) could also be significant given the complex influencing mechanisms. This study could contribute to the existing literature with statistically significant paths and established validity and reliability. Not all links between the proposed variables were tested in this study due to our theoretical and practical consideration of the existing literature, which influenced the explanatory power. Future researchers may further examine the interrelationships underexplored among our proposed variables. Second, this study focused on higher education students, while the variables might be extended to other educational levels and technologies for a more comprehensive understanding. Third, the demographic information lacked some details, such as the affiliations of our participants and majors. However, the representativeness of our collected dataset could be justified by other aspects, such as our data collection sources and the identified provinces and regions of the participants. Future studies should be encouraged to collect a more comprehensive and balanced population in genders, ages, majors, and areas.

Implications for future research. Our findings have both theoretical and practical implications. Theoretically, this study incorporates new external variables to UTAUT, including trust in feedback, peer influence, interactivity, personal investment, willingness for e-learning, and instructional support. These features can reveal how specific perceptual and systemic factors influence higher education students' acceptance and adoption of Grammarly. This study also extends the traditional UTAUT to Grammarly utilization, bridging the scarcity of the existing literature. The results can enrich technology acceptance theories that are conceptually similar or related to UTAUT and offer evidence regarding their adaptivity to emerging contexts. This study adds a specific example of automated writing evaluation tool utilization, which aligns with the implications of Ding and Zou's systematic review (2024). This study clarifies higher education students' psychological mechanisms interacting with automated writing evaluation tools. For example, the role of perceived interactivity established in this study further strengthens the significance of interactive learning environments (e.g., Chang et al. 2023).

This study sets an example of exploring innovative and detailed variables in technology acceptance research. Future researchers are encouraged to validate the specific benefits of automated writing evaluation tools in technology acceptance models at different educational levels, writing tasks, and linguistic levels (such as vocabulary, grammatical correctness, and content organization) (e.g., McCarthy et al. 2022). Such methods will allow technology acceptance research to further contribute to domain-specific language education research, congruent with a sustained tendency to apply objective approaches to language testing. Self-designed and adapted measurements validated in this study can provide references to extend our findings to language learners and assessments at various educational levels, presumably in association with other innovative variables. To generate innovative external predictors of technology acceptance and adoption, Yu et al. (2024) study sequentially combined a grounded theory approach, fuzzy set qualitative comparative analysis, and structural equation modeling techniques. Moreover, keyword and publication distribution analysis through bibliographic techniques can also enlighten educational technology researchers, for instance, to dig out innovative predictors and effectiveness indicators (Wang et al. 2024; Lin and Yu 2024b).

Practically, teachers need to guide students in using automated writing evaluation tools like Grammarly so that more students may benefit from it. With the popularity of computer-mediated collaborative learning in teaching writing skills, frameworks and theories capturing teachers' and students' feedback literacy should be explored in the applications of automated writing evaluation (e.g., Saeed and Alharbi 2023). This study offers evidence regarding the quality of automated writing feedback by examining students' trust. Nevertheless, feedback literacy encompasses learners' openness to feedback, engagement with feedback, and the enactment of feedback (Woitt et al. 2023). Other researchers framed feedback literacy to include information-seeking, sense-making, feedback utilization, and emotional regulation in dealing with feedback (Dawson et al. 2023). This study only adds limited contributions to the literacy of automated feedback, especially in the present time of artificial intelligence-empowered writing assessment. Future studies may further explore feedback-seeking behavior with automated writing evaluation tools to enrich the conceptualizations and theories of feedback literacy. For example, investigations can enrich the understanding and pedagogical implications of using automated writing evaluation to assist peer assessment activity and enhance teachers' feedback literacy (e.g., Li 2023; Link et al. 2022).

Proper instruction may enlarge the positive impact of peer modeling on students' acceptance and use of automated writing evaluation tools. The unstable impact of peer influence should arouse educators' awareness of students' peer pressure when they adopt educational technologies. Successful transformation to digitalized learning and workflow needs proper guidance to enhance students' willingness. As for specific approaches, proper instruction can be strengthened by teachers, technology designers, service providers, and other resources. Given the complementary role of automated writing evaluation in some studies, future instructors may consider combining teacher, peer, and automated writing feedback to enhance student engagement and quality of writing (e.g., Zhang and Hyland 2022). Future automated writing evaluation designs may attach significance to interactive features, accuracy, reliability, and instruction on use to enhance students' acceptance and use of automated writing evaluation tools at all educational levels.

Data availability

The datasets collected and analyzed during the current study should be retrieved upon reasonable requests from the corresponding author based on appropriate reasons. The data availability statement of this article is open in the OSF respiratory at https://osf.io/gbtqe/?view_only=38210688ba1048eca93427acbc555daa.

Received: 9 March 2024; Accepted: 13 September 2024;

Published online: 18 August 2025

References

- Abbad M (2021) Using the UTAUT model to understand students' usage of e-learning systems in developing countries. *Educ Inf Technol* 26:7205–7224. <https://doi.org/10.1007/s10639-021-10573-5>
- Aboagye E, Yawson JA, Appiah KN (2021) COVID-19 and e-learning: The challenges of students in tertiary institutions. *Soc Educ Res* 1–8. <https://doi.org/10.37256/ser.212021422>
- Abubakir H, Alshaboul Y (2023) Unravelling EFL teachers' mastery of TPACK: Technological pedagogical and content knowledge in writing classes. *Heliyon* 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17348>
- Al-Adwan AS, Yaseen H, Alsoud A, Abousweilem F, Al-Rahmi WM (2022) Novel extension of the UTAUT model to understand continued usage intention of

- learning management systems: The role of learning tradition. *Educ Inf Technol* 27:3567–3593. <https://doi.org/10.1007/s10639-021-10758-y>
- Al-Ghaith W (2015) Applying the technology acceptance model to understand social networking sites (SNS) usage: Impact of perceived social capital. *Int J Comput Sci Inf Technol* 7(4):105–117. <https://doi.org/10.5121/ijcsit.2015.7409>
- AL-Sayid F, Kirkil G (2023) Exploring non-linear relationships between perceived interactivity or interface design and acceptance of collaborative web-based learning. *Educ Inf Technol* 1–48. <https://doi.org/10.1007/s10639-023-11635-6>
- Berhan Y, Dessalegn M, Hetey A, Assamnew B, Tsegaye S, Makonnen M, Abebe S (2023) Learners and system readiness for digital learning in the Ethiopian health sector: the path to blended learning. *Int J Mob Learn Organ* 17(4):557–573. <https://doi.org/10.1504/IJMLIO.2023.133773>
- Bhattacharjee A (2001) Understanding information systems continuance: An expectation-confirmation model. *MIS Quart* 351–370. <https://doi.org/10.2307/3250921>
- Bond M, Marin VI, Dolch C, Bedenlier S, Zawacki-Richter O (2018) Digital transformation in German higher education: student and teacher perceptions and usage of digital media. *Int J Educ Technol High Educ* 15(1):48. <https://doi.org/10.1186/s41239-018-0130-1>
- Budhathoki T, Zirar A, Njoya ET, Timsina A (2024) ChatGPT adoption and anxiety: a cross-country analysis utilising the unified theory of acceptance and use of technology (UTAUT). *Stud Higher Educ* 1–16. <https://doi.org/10.1080/03075079.2024.2333937>
- Chang A (2012) UTAUT and UTAUT 2: A review and agenda for future research. *Winners* 13(2):106–114
- Chang CW, Heo J, Yeh SC, Han HY, Li M (2018) The effects of immersion and interactivity on college students' acceptance of a novel VR-supported educational technology for mental rotation. *IEEE Access* 6:66590–66599. <https://doi.org/10.1109/ACCESS.2018.2878270>
- Chang HH, Fu CS, Huang CY (2017) Willingness to adopt or reuse an e-learning system: The perspectives of self-determination and perceived characteristics of innovation. *Innov Educ Teach Int* 54(5):511–520. <http://doi-org.s.weebvnp.zju.edu.cn:8001/10.1080/14703297.2016.1194768>
- Chang TS, Li Y, Huang HW, Whitfield B (2021) Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback. In *Proceedings of the 2021 2nd International Conference on Education Development and Studies* (pp. 31–35). <https://doi.org/10.1145/3459043.3459065>
- Chang YS, Chou CH, Chuang MJ, Li WH, Tsai IF (2023) Effects of virtual reality on creative design performance and creative experiential learning. *Interact Learn Environ* 31(2):1142–1157. <https://doi.org/10.1080/10494820.2020.1821717>
- Chatterjee S, Bhattacharjee KK, Tsai CW, Agrawal AK (2021) Impact of peer influence and government support for successful adoption of technology for vocational education: A quantitative study using PLS-SEM technique. *Qual Quant* 55:2041–2064. <https://doi.org/10.1007/s11135-021-01100-2>
- Cohen J (1988) Statistical power analysis for the behavioral sciences. In *Erlbaum association* (2nd ed.). Lawrence Erlbaum Associates
- Costley J, Lange C (2023) The mediating effect of sequencing and fading on the relationship between e-learning personalization and germane load. *Interact Learn Environ* 31(6):3436–3446. <https://doi.org/10.1080/10494820.2021.1929334>
- Dautzenberg PS, Voß GM (2022) A theoretical framework for trust in automation considering its relationship to technology acceptance and its influencing factors. In *AHFE International Conference* (Vol. 27). <https://doi.org/10.54941/ahfe1002462>
- Dawson P, Yan Z, Lipnevich A, Tai J, Boud D, Mahoney P (2023) Measuring what learners do in feedback: The feedback literacy behaviour scale. *Assessment & Evaluat Higher Educ* 1–15. <https://doi.org/10.1080/02602938.2023.2240983>
- Deeva G, Bogdanova D, Serral E, Snoeck M, De Weerd J (2021) A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Comput Educ* 162:104094. <https://doi.org/10.1016/j.compedu.2020.104094>
- Deng X, Yu Z (2023) An extended hedonic motivation adoption model of TikTok in higher education. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-023-11749-x>
- Ding L, Zou D (2024) Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Educ Inf Technol* 1–53. <https://doi.org/10.1007/s10639-023-12402-3>
- Dörnyei Z (2007) *Research methods in applied linguistics*. Oxford University Press
- Fan N (2023) Exploring the effects of automated written corrective feedback on EFL students' writing quality: A mixed-methods study. *SAGE Open* 13(2):21582440231181296. <https://doi.org/10.1177/21582440231181296>
- Farooq MS, Salam M, Jaafar N, Fayolle A, Ayupp K, Radovic-Markovic M, Sajid A (2017) Acceptance and use of lecture capture system (LCS) in executive business studies: Extending UTAUT2. *Interact Technol Smart Educ* 14(4):329–348. <https://doi.org/10.1108/ITSE-06-2016-0015>
- Fritia TN (2021) Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: J Engl Lang. Lit, Teach* 5(1):65–78. <https://doi.org/10.31002/metathesis.v5i1.3519>
- Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 18(1):39–50. <https://doi.org/10.1177/002224378101800104>
- Foster S (2019) What barriers do students perceive to engagement with automated immediate formative feedback. *J Interactive Media Educ* 2019(1). <https://doi.org/10.5334/jime.516>
- Fu QK, Zou D, Xie H, Cheng G (2022) A review of AWE feedback: types, learning outcomes, and implications. *Comput Assisted Lang Learn* 1–43. <https://doi.org/10.1080/09588221.2022.2033787>
- Gao J (2021) Exploring the feedback quality of an automated writing evaluation system Pigai. *Int J Emerg Technol Learn* 16(11):322–330. <https://doi.org/10.3991/ijet.v16i11.19657>
- Gao Z, Cheah JH, Lim XJ, Luo X (2024) Enhancing academic performance of business students using generative AI: An interactive-constructive-active-passive (ICAP) self-determination perspective. *Int J Manag Educ* 22(2):100958. <https://doi.org/10.1016/j.jime.2024.100958>
- Gedrimiene E, Celik I, Kaasila A, Mäkitalo K, Muukkonen H (2024) Artificial intelligence (AI)-enhanced learning analytics (LA) for supporting career decisions: Advantages and challenges from user perspective. *Educ Inf Technol* 29(1):297–322. <https://doi.org/10.1007/s10639-023-12277-4>
- Geng J, Razali AB, Ismail L, Noordin N (2024) Effectiveness of automated writing evaluation program to enhance EFL undergraduates' writing performance in China. *Asia Pacific Educ Rev* 1–16. <https://doi.org/10.1007/s12564-023-09922-5>
- Go H, Kang M, Suh SC (2020) Machine learning of robots in tourism and hospitality: interactive technology acceptance model (iTAM)—cutting edge. *Tour Rev* 75(4):625–636. <https://doi.org/10.1108/TR-02-2019-0062>
- Grammarly (2024) Retrieved from <https://www.grammarly.com/>
- Guo Q, Feng R, Hua Y (2022) How effectively can EFL students use automated written corrective feedback (AWCF) in research writing? *Comput Assist Lang Learn* 35(9):2312–2331. <https://doi.org/10.1080/09588221.2021.1879161>
- Hair JF, Hult GTM, Ringle CM (2014) *Partial least squares structural equation modeling (PLS-SEM)*. SAGE Publications
- Hair JF, Risher JJ, Sarstedt M, Ringle CM (2019) When to use and how to report the results of PLS-SEM. *Eur Bus Rev* 31(1):2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair Jr, JF, Hult GTM, Ringle CM, Sarstedt M (2021) *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications
- Henseler J, Ringle CM, Sarstedt M (2014) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci* 43(1):115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Herting DC, Cladellas Pros R, Castelló Tarrida A (2023) Habit and social influence as determinants of PowerPoint use in higher education: A study from a technology acceptance approach. *Interact Learn Environ* 31(1):497–513. <https://doi.org/10.1080/10494820.2020.1799021>
- Hidayat-ur-Rehman I, Ibrahim Y (2023) Exploring factors influencing educators' adoption of ChatGPT: a mixed method approach. *Interactive Technol Smart Educ* 1–36. <https://doi.org/10.1108/ITSE-07-2023-0127>
- Horta H, Li H (2023) Nothing but publishing: The overriding goal of PhD students in mainland China, Hong Kong, and Macau. *Stud High Educ* 48(2):263–282. <https://doi.org/10.1080/03075079.2022.2131764>
- Hou YK, Yu ZG (2023) The unified theory of acceptance and use of DingTalk for educational purposes in China: An extended structural equation model. *Humanities Soc Sci Commun* 10(1):1–12. <https://doi.org/10.1057/s41599-023-02257-x>
- Huang X, Xu W, Li F, Yu Z (2024) A meta-analysis of effects of automated writing evaluation on anxiety, motivation, and second language writing skills. *Asia-Pac Educ Res* 33:957–976. <https://doi.org/10.1007/s40299-024-00865-y>
- Huang Y, Xu W, Sukjairungwattana P, Yu Z (2024) Learners' continuance intention in multimodal language learning education: An innovative multiple linear regression model. *Heliyon* 10(6):E28104. <https://doi.org/10.1016/j.heliyon.2024.e28104>
- Huawei S, Aryadoust V (2023) A systematic review of automated writing evaluation systems. *Educ Inf Technol* 28(1):771–795. <https://doi.org/10.1007/s10639-022-11260-9>
- Hutcheson GD, Sofroniou N (1999) *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage Publications Ltd
- Hwang GJ, Chen PY (2023) Effects of a collective problem-solving promotion-based flipped classroom on students' learning performances and interactive patterns. *Interact Learn Environ* 31(5):2513–2528. <https://doi.org/10.1080/10494820.2019.1568263>
- Ismatullaev UVU, Kim SH (2024) Review of the factors affecting acceptance of AI-infused systems. *Hum Factors* 66(1):126–144. <https://doi.org/10.1177/00187208211064707>

- Johnson JB, Reddy P, Chand R, Naiker M (2021) Attitudes and awareness of regional Pacific Island students towards e-learning. *Int J Educ Technol High Educ* 18(1):1–20. <https://doi.org/10.1186/s41239-021-00248-z>
- Kassim ES, Jailani SFAK, Hairuddin H, Zamzuri NH (2012) Information system acceptance and user satisfaction: The mediating role of trust. *Procedia-Soc Behav Sci* 57:412–418. <https://doi.org/10.1016/j.sbspro.2012.09.1205>
- Kennewell S, Tanner H, Jones S, Beauchamp G (2008) Analysing the use of interactive technology to implement interactive teaching. *J Comput Assist Learn* 24(1):61–73. <https://doi.org/10.1111/j.1365-2729.2007.00244.x>
- Kikawa CR, Kiconco C, Agaba M, Ntirampaba D, Ssematimba A, Kalema BM (2022) Social media marketing for small and medium enterprise performance in Uganda: A structural equation model. *Sustainability* 14(21):14391. <https://doi.org/10.3390/su142114391>
- Kim HW, Chan HC, Gupta S (2007) Value-based adoption of mobile internet: an empirical investigation. *Decis Support Syst* 43(1):111–126. <https://doi.org/10.1016/j.dss.2005.05.009>
- King RB, Yeung SSS, Cai Y (2019) Personal investment theory: A multi-faceted framework to understand second and foreign language motivation. *System* 86:102123. <https://doi.org/10.1016/j.system.2019.102123>
- Kline RB (2015) In DA Kenny & TD Little (Eds.), *Principles and practice of structural equation modeling*. Guilford publications
- Kloppers J (2023) Errors in errors: an exploration of Grammarly's corrective feedback. *Int J Comput-Assist Lang Learn Teach* 13(1):1–16. <https://doi.org/10.4018/IJCALLT.325792>
- Kock N (2015) Common method bias in PLS-SEM: A full collinearity assessment approach. *Int J e-Collaboration* 11(4):1–10. <https://doi.org/10.4018/ijec.2015100101>
- Kolb DA (2014) *Experiential learning: Experience as the source of learning and development*. FT press
- Koltovskaia S (2020) Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assess Writ* 44:100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Lakhal S, Khechine H (2021) Technological factors of students' persistence in online courses in higher education: The moderating role of gender, age and prior online course experience. *Educ Inf Technol* 26:3347–3373. <https://doi.org/10.1007/s10639-020-10407-w>
- Li AW (2023) Using Perceptiv to support AI-based online writing assessment across the disciplines. *Assess Writ* 57:100746. <https://doi.org/10.1016/j.asw.2023.100746>
- Li M, Wang T (2024) Optimizing learning return on investment: Identifying learning strategies based on user behavior characteristic in language learning applications. *Educ Inf Technol* 29:6651–6681. <https://doi.org/10.1007/s10639-023-12078-9>
- Li R (2021) Modeling the continuance intention to use automated writing evaluation among Chinese EFL learners. *SAGE Open* 11(4):21582440211060782. <https://doi.org/10.1177/21582440211060782>
- Li R, Meng Z, Tian M, Zhang Z, Ni C, Xiao W (2019) Examining EFL learners' individual antecedents on the adoption of automated writing evaluation in China. *Comput Assist Lang Learn* 32(7):784–804. <https://doi.org/10.1080/09588221.2018.1540433>
- Lim K, Song J, Park J (2023) Neural automated writing evaluation for Korean L2 writing. *Nat Lang Eng* 29(5):1341–1363. <https://doi.org/10.1017/S1351324922000298>
- Lin Y, Yu Z (2023) Extending technology acceptance model to higher-education students' use of digital academic reading tools on computers. *Int J Educ Technol High Educ* 20:34. <https://doi.org/10.1186/s41239-023-00403-8>
- Lin Y, Yu Z (2024a) Learner perceptions of artificial intelligence-generated pedagogical agents in language learning videos: Embodiment effects on technology acceptance. *Int J Human-Comput Interact* 1-22. <https://doi.org/10.1080/10447318.2024.2359222>
- Lin Y, Yu Z (2024b) An integrated bibliometric analysis and systematic review modelling students' technostress in higher education. *Behav Inf Technol* 1-25. <https://doi.org/10.1080/0144929X.2024.2332458>
- Link S, Mehrzad M, Rahimi M (2022) Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Comput Assist Lang Learn* 35(4):605–634. <https://doi.org/10.1080/09588221.2020.1743323>
- Liu C, Hou J, Tu YF, Wang Y, Hwang GJ (2021) Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interact Learning Environ* 1-19. <https://doi.org/10.1080/10494820.2021.2012812>
- Locke E, Latham G (2015) Goal-setting theory. In *Organizational Behavior* 1 (pp. 159–183). Routledge
- Lowry PB, Gaskin J (2014) Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Trans Professional Commun* 57(2):123–146. <https://doi.org/10.1109/TPC.2014.2312452>
- Mai DTT, Da CV, Hanh NV (2024) The use of ChatGPT in teaching and learning: A systematic review through SWOT analysis approach. *Front Educ* 9:1328769. <https://doi.org/10.3389/feeduc.2024.1328769>
- McCarthy KS, Roscoe RD, Allen LK, Likens AD, McNamara DS (2022) Automated writing evaluation: Does spelling and grammar feedback support high-quality writing and revision? *Assess Writ* 52:100608. <https://doi.org/10.1016/j.asw.2022.100608>
- Meschitti V (2019) Can peer learning support doctoral education? Evidence from an ethnography of a research team. *Stud High Educ* 44(7):1209–1221. <https://doi.org/10.1080/03075079.2018.1427711>
- Mishra A, Shukla A, Rana NP, Currie WL, Dwivedi YK (2023) Re-examining post-acceptance model of information systems continuance: a revised theoretical model using MASEM approach. *Int J Inf Manag* 68:102571. <https://doi.org/10.1016/j.jinfomgt.2022.102571>
- Mou J, Shin DH, Cohen J (2017) Understanding trust and perceived usefulness in the consumer acceptance of an e-service: a longitudinal investigation. *Behav Inf Technol* 36(2):125–139. <https://doi.org/10.1080/0144929X.2016.1203024>
- Nguyen BTT (2023) Using online tools to help students elevate lexical range in English-as-a-foreign-language writing. *RELC J* 00336882231217528. <https://doi.org/10.1177/00336882231217528>
- Nunes A, Cordeiro C, Limpo T, Castro SL (2022) Effectiveness of automated writing evaluation systems in school settings: A systematic review of studies from 2000 to 2020. *J Comput Assist Learn* 38(2):599–620. <https://doi.org/10.1111/jcal.12635>
- Obeid A, Ibrahim R, Fadhil A (2024) Extended model of expectation confirmation model to examine users' continuous intention toward the utilization of e-learning platforms. *IEEE Access* 12:40752–40764. <https://doi.org/10.1109/ACCESS.2024.3373190>
- Obilor EI (2023) Convenience and purposive sampling techniques: Are they the same. *Int J Innovat Soc Sci Educ Res* 11(1):1–7
- Ozili PK (2023) The acceptable R-square in empirical modelling for social science research. In *Social research methodology and publishing results: A guide to non-native English speakers* (pp. 134–143). IGI Global
- Papakostas C, Troussas C, Krouska A, Sgouropoulou C (2023) Exploring users' behavioral intention to adopt mobile augmented reality in education through an extended technology acceptance model. *Int J Hum -Comput Interact* 39(6):1294–1302. <https://doi.org/10.1080/10447318.2022.2062551>
- Pasupuleti RS, Thiyyagura D (2024) An empirical evidence on the continuance and recommendation intention of ChatGPT among higher education students in India: An extended technology continuance theory. *Educ Inf Technol* 1-21. <https://doi.org/10.1007/s10639-024-12573-7>
- Patil H, Undale S (2023) Willingness of university students to continue using e-learning platforms after compelled adoption of technology: Test of an extended UTAUT model. *Educ Inf Technol* 28(11):14943–14965. <https://doi.org/10.1007/s10639-023-11778-6>
- Pham L, Limbu YB, Bui TK, Nguyen HT, Pham HT (2019) Does e-learning service quality influence e-learning student satisfaction and loyalty? Evidence from Vietnam. *Int J Educ Technol High Educ* 16(1):1–26. <https://doi.org/10.1186/s41239-019-0136-3>
- Raffaghelli JE, Rodríguez ME, Guerrero-Roldán AE, Bañeres D (2022) Applying the UTAUT model to explain the students' acceptance of an early warning system in higher education. *Comput Educ* 182:104468. <https://doi.org/10.1016/j.compedu.2022.104468>
- Rafiee M, Abbasian-Naghnesh S (2021) E-learning: development of a model to assess the acceptance and readiness of technology among language learners. *Comput Assist Lang Learn* 34(5-6):730–750. <https://doi.org/10.1080/09588221.2019.1640255>
- Ranalli J (2021) L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *J Second Lang Writ* 52:100816. <https://doi.org/10.1016/j.jslw.2021.100816>
- Rasheed RA, Kamsin A, Abdullah NA (2020) Challenges in the online component of blended learning: A systematic review. *Comput Educ* 144:103701. <https://doi.org/10.1016/j.compedu.2019.103701>
- Raza SA, Qazi W, Khan KA, Salam J (2020) Social isolation and acceptance of the learning management system (LMS) in the time of COVID-19 pandemic: An expansion of the UTAUT model. *J Educ Comput Res* 59(2):183–208. <https://doi.org/10.1177/0735633120960421>
- Ringle CM, Wende S, Becker JM (2024) *SmartPLS 4*. Bönningstedt: SmartPLS. Retrieved from <https://www.smartpls.com>
- Roscoe RD, Wilson J, Johnson AC, Mayra CR (2017) Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Comput Hum Behav* 70:207–221. <https://doi.org/10.1016/j.chb.2016.12.076>
- Rotsaert T, Panadero E, Schellens T, Raes A (2018) "Now you know what you're doing right and wrong!" Peer feedback quality in synchronous peer assessment in secondary education. *Eur J Psychol Educ* 33:255–275. <https://doi.org/10.1007/s10212-017-0329-x>
- Saeed MA, Alharbi MA (2023) Towards fostering Saudi EFL learners' collaborative engagement and feedback literacy in writing. *Assess Writ* 56:100721. <https://doi.org/10.1016/j.asw.2023.100721>

- Saricaoglu A, Bilki Z (2021) Voluntary use of automated writing evaluation by content course students. *ReCALL* 33(3):265–277. <https://doi.org/10.1017/S0958344021000021>
- Sarstedt M, Mooi E (2014) A concise guide to market research: The process, data, and, methods using IBM SPSS statistics (2nd edition). London: Springer
- Schworm S, Gruber H (2012) E-learning in universities: Supporting help-seeking processes by instructional prompts. *Br J Educ Technol* 43(2):272–281. <https://doi.org/10.1111/j.1467-8535.2011.01176.x>
- Strzelecki A (2024) Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innov High Educ* 49:223–245. <https://doi.org/10.1007/s10755-023-09686-1>
- Sun H, Guo R (2024) The effect of metaverse concert characteristics on user re-participation intention: A combination of the ETAM and the SPICE model. *Int J Human-Comput Interaction*, 1–14. <https://doi.org/10.1080/10447318.2024.2331859>
- Sun Z, Yang Y (2023) The mediating role of learner empowerment in the relationship between the community of inquiry and online learning outcomes. *Internet High Educ* 58:100911. <https://doi.org/10.1016/j.iheduc.2023.100911>
- Sungur-Gül K, Ateş H (2021) Understanding pre-service teachers' mobile learning readiness using theory of planned behavior. *Educ Technol Soc* 24(2):44–57. <https://www.jstor.org/stable/27004930>
- Tian L, Zhou Y (2020) Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System* 91:102247. <https://doi.org/10.1016/j.system.2020.102247>
- Trivedi SK, Patra P, Singh S (2022) A study on intention to use social media in higher education: the mediating effect of peer influence. *Glob Knowl, Mem Commun* 71(1/2):52–69. <https://doi.org/10.1108/GKMC-11-2020-0169>
- van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84(2):523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Venkatesh V, Brown SA, Maruping LM, Bala H (2008) Predicting different conceptualizations of system use: The competing roles of behavioral intention, facilitating conditions, and behavioral expectation. *MIS Q* 32(3):483–502. <https://doi.org/10.2307/25148853>
- Venkatesh V, Morris MG (2000) Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quart* 115–139. <https://doi.org/10.2307/3250981>
- Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: Toward a unified view. *MIS Q* 27(3):425–478. <https://doi.org/10.2307/30036540>
- Venkatesh V, Thong JY, Chan FK, Hu PJH, Brown SA (2011) Extending the two-stage information systems continuance model: Incorporating UTAUT predictors and the role of context. *Inf Syst J* 21(6):527–555. <https://doi.org/10.1111/j.1365-2575.2011.00373.x>
- Vorm ES, Combs DJ (2022) Integrating transparency, trust, and acceptance: The intelligent systems technology acceptance model (ISTAM). *Int J Hum –Comput Interact* 38(18–20):1828–1845. <https://doi.org/10.1080/10447318.2022.2070107>
- Wang CL, Chen XJ, Yu T, Liu YD, Jing YH (2024) Education reform and change driven by digital technology: A bibliometric study from a global perspective. *Human Soc Sci Commun* 11:256. <https://doi.org/10.1057/s41599-024-02717-y>
- Wang EL, Matsumura LC, Correnti R, Litman D, Zhang H, Howe E, Quintana R (2020) eRevis (ing): Students' revision of text evidence use in an automated writing evaluation system. *Assess Writ* 44:100449. <https://doi.org/10.1016/j.asw.2020.100449>
- Wang Y, Luo X, Liu CC, Tu YF, Wang N (2022) An integrated automatic writing evaluation and SVVR approach to improve students' EFL writing performance. *Sustainability* 14(18):11586. <https://doi.org/10.3390/su141811586>
- Wang Y, Yu L, Yu ZG (2022) An extended CCTalk technology acceptance model in EFL education. *Educ Inf Technol* 27(5):6621–6640. <https://doi.org/10.1007/s10639-022-10909-9>
- Wei P, Wang X, Dong H (2023) The impact of Automated writing evaluation (AWE) on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Front Psychol* 14:1249991. <https://doi.org/10.3389/fpsyg.2023.1249991>
- Woitt S, Weidlich J, Jivet I, Orhan Gökşin D, Drachsler H, Kalz M (2023) Students' feedback literacy in higher education: An initial scale validation study. *Teaching Higher Educ* 1–20. <https://doi.org/10.1080/13562517.2023.2263838>
- Wu JH, Tennyson RD, Hsia TL (2010) A study of student satisfaction in a blended e-learning system environment. *Comput Educ* 55(1):155–164. <https://doi.org/10.1016/j.compedu.2009.12.012>
- Yang H, Gao C, Shen HZ (2023) Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Educ Inf Technol*, 1–22. <https://doi.org/10.1007/s10639-023-11991-3>
- Yao Y, Wang W, Yang X (2021) Perceptions of the inclusion of automatic writing evaluation in peer assessment on EFL writers' language mindsets and motivation: A short-term longitudinal study. *Assess Writ* 50:100568. <https://doi.org/10.1016/j.asw.2021.100568>
- Yu L, Chen Z, Yao P, Liu H (2021) A study on the factors influencing users' online knowledge paying-behavior based on the UTAUT model. *J Theor Appl Electron Commer Res* 16(5):1768–1790. <https://doi.org/10.3390/jtaer16050099>
- Yu T, Dai J, Chen X, Wang C (2024) Factors influencing continuance intention in blended learning among business school students in China: based on grounded theory and FsQCA. *Interactive Learn Environ* 1–28. <https://doi.org/10.1080/10494820.2024.2370477>
- Yu Z, Yu L (2023) Examining factors that influence learner retention in MOOCs during the COVID-19 pandemic time. *Sage Open* 13(2):21582440231175371. <https://doi.org/10.1177/21582440231175371>
- Yu Z, Yu X (2019) An extended technology acceptance model of a mobile learning technology. *Comput Appl Eng Educ* 27(3):721–732. <https://doi.org/10.1002/cae.22111>
- Zadorozhnyy A, Lee JS (2023) Informal Digital Learning of English and willingness to communicate in a second language: self-efficacy beliefs as a mediator. *Comput Assisted Lang Learn* 1–21. <https://doi.org/10.1080/09588221.2023.2215279>
- Zhai N, Ma X (2022) Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Comput Assist Lang Learn* 35(9):2817–2842. <https://doi.org/10.1080/09588221.2021.1897019>
- Zhai N, Ma X (2023) The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *J Educ Comput Res* 61(4):875–900. <https://doi.org/10.1177/07356331221127300>
- Zhang Y, Liu G (2022) Revisiting informal digital learning of English (IDLE): A structural equation modeling approach in a university EFL context. *Comput Assisted Lang Learn*. <https://doi.org/10.1080/09588221.2022.2134424>
- Zhang Z, Hyland K (2022) Fostering student engagement with feedback: An integrated approach. *Assess Writ* 51:100586. <https://doi.org/10.1016/j.asw.2021.100586>

Acknowledgements

The authors would like to extend sincere gratitude to their funders, journal editors, anonymous reviewers, and all other participants who offered help with reviewing and publishing this paper. The authors sincerely thank Professor Peng Duan for his generous and insightful suggestions during our revisions. The first author would like to express his special gratitude to his supervisors for the current program at Beijing Language and Culture University and to Ms. Daojing Li, a postgraduate in applied foreign linguistics at Beijing Language and Culture University, for her contributions to reviewing automated writing evaluation tools in an earlier draft of this article. This work is supported by Key Research and Application Project of the Key Laboratory of Key Technologies for Localization Language Services of the State Administration of Press and Publication, "Research on Localization and Intelligent Language Education Technology for the 'Belt and Road Initiative'" [CSLS 20230012], Special fund of Beijing Co-construction Project-Research and reform of the "Undergraduate Teaching Reform and Innovation Project" of Beijing higher education in 2020-innovative "multilingual +" excellent talent training system [202010032003], and a BLCU supported project for young researchers program (supported by the Fundamental Research Funds for the Central Universities) [24YCX096].

Author contributions

YL (First author): Conceptualization; Investigation; Data curation; Formal analysis; Visualization; Project Administration; Writing—original draft; Writing—review and editing; Funding acquisition. ZY (Corresponding author): Funding acquisition.

Competing interests

The authors declare no competing interests. Particularly, the authors do not represent any interest of the Grammarly Inc. or compete with it.

Ethical approval

This study was ethically approved by the Academic Committee of Beijing Language and Culture University.

Informed consent

Informed consent was obtained from all participants and/or their legal guardians for participation in the study. All responses collected are confidential and only used for this study.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-03861-1>.

Correspondence and requests for materials should be addressed to Zhonggen Yu.

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