

ChatGPT as a Feedback Provider: Arab Students' Perceptions of AI-Based Writing Evaluation in Academic English

Received 2025-11-11
Accepted 2026-05-19
Published 2026-07-04

Abdullah Alshayban

Department of English Language and Literature, Qassim University,
Qassim 52571, Saudi Arabia
ashieban@qu.edu.sa

To cite this article: Alshayban, Abdullah. (2026). ChatGPT as a Feedback Provider: Arab Students' Perceptions of AI-Based Writing Evaluation in Academic English. *Ijaz Arabi: Journal of Arabic Learning*, 9 (3), 1082-1098, DOI: <https://doi.org/10.18860 /ijazarabi.V9i3.42134>

Abstract

The research has four constructs, including perceived usefulness, trust and accuracy, motivation, and behavior intention, which indicate the perceptions of EFL students regarding ChatGPT as a feedback tool in English writing. The mixed-methods design was used to collect data (N = 254) on surveys and open-ended responses. Quantitative methods (descriptive statistics, reliability, regression, and mediation) indicated that ChatGPT was the most useful tool overall for enhancing readers' grammar, organization, and clarity. Its performance was impressive, but its accuracy was inconsistent, with occasional gaps in context. The regression showed that perceived usefulness, trust, and motivation were all useful predictors of behavioral intention, and that motivation mediated the effect of usability on behavioral intention. This was supported by the qualitative results, which focused on ChatGPT's immediacy and supportiveness. ChatGPT is a useful, supplementary source of feedback compared to a teacher's because it helps foster a student's self-confidence and self-reliance when editing drafts.

Keywords: ChatGPT; feedback; human vs. AI; EFL students, motivation; perceived usefulness; trust and accuracy; behavioral intention

INTRODUCTION

Feedback is a crucial component of learning L2 (second language) or EFL (English as a Foreign Language), as it assesses learners' lagging areas and assists them in addressing their weaknesses. Just as the world is being pushed toward artificial intelligence, so are the feedback platforms. Moreover, the traditional feedback channels, whether teacher-based or peer-based, consume time, lack uniformity, contain bias, and make it hard for students to reach out for feedback. To overcome these issues, students are now being pointed toward AI-generated feedback because it provides contextualized, hassle-free, and immediate feedback, minimizing biased and judgmental remarks.

ChatGPT is quite eminent among students for providing feedback. Thus, a plethora of researchers are trying to measure the effectiveness of this tool and to determine how it can be used to improve learners' EFL writing skills. Apparently, AI-based feedback increases not just surface-level accuracy (e.g., grammar and vocabulary) but also the entire extent of discourse organization and linkage. It also promotes a higher degree of autonomy and self-regulation of learners (Polakova & Ivenz, 2024). Also, Kurt and Kurt (2024) demonstrated that ChatGPT, when positioned as an automated feedback tool, can provide iterative and individualized support that approximates the scaffolding of a human tutor. Combined, these studies give a clue on the potential of AI feedback to

solve the issues associated with conventional types of feedback to expand access and opportunities in the classrooms teaching EFL writing.

Comparisons of AI-generated and human feedback, however, show that there is complexity in the learners, as well as accuracy and trust. Even though students enjoyed the immediacy and clarity of the feedback that ChatGPT provides, Alnemrat (2025) reported that the feedback provided by teachers is more likely to be trusted and relevant to the context. This implies that learners are cautious about trusting the comments that AI systems generate and have a certain degree of optimism regarding the benefits of automation, while still maintaining a level of doubt about the credibility of the comments. Escalante (2023), in a study of English as a New Language (ENL) learners, also found that although AI feedback was part of students' measurable linguistic progress, students expressed ambivalence about relying too heavily on automated systems due to concerns about thinking critically and being original in writing. Taken together, these findings seem to indicate that while learner perceptions are not monolithic, the perception of usefulness, trust, and motivational impact interact with each other.

Besides the aspect of trust and accuracy, the true motivation of learners to use ChatGPT is beginning to be recognized. In one class ChatGPT was assigned as an L2 writing assistant, and, due to its dialogic, supportive, and nonjudgmental nature, it was helpful for revising and playing (Mizumoto et al., 2024). The authors explained that synchronous feedback was preferable for students, as there was no delay involved in receiving a teacher's feedback with synchronous feedback. On the other hand, they were also concerned about the possibility of "deskilling," as students could blindly follow the AI-generated feedback and fail to use their cognitive skills and analytical thinking (Mizumoto et al., 2024). This dual effect raises an interesting aspect in further research about the role of motivation as a mediator for the uptake of feedback and in which circumstances AI feedback reinforces sustainable learning behaviors.

New research also suggests that people interpret AI feedback differently depending on context. According to Nelson et al. (2025), feelings of both optimism and caution are evident among students in higher education institutions in Latin America. The range of supporters experienced increased confidence and motivation, yet others raised issues of plagiarism, ethics, and homogenization of writing style (Nelson et al., 2025). This study asserts that generative AI feedback cannot be regarded as the same as human feedback due to the context of cultural, ethical, and pedagogical perspectives. Collectively, these results indicate that perceived usefulness, trust/accuracy, motivation, and behavioral intention to continue to use ChatGPT should all be considered when investigating student perceptions regarding ChatGPT.

One of the key factors that influences learners' adoption of AI-based feedback on academic writing is trust. Zhang et al. (2025) reported a comparison of student trust for AI and for human and co-produced feedback: students described ChatGPT as an immediate and extensive list of comments but retained skepticism about the contextual correctness of AI-generated comments. Although some AI feedback might appear linguistically correct, it does not provide the quality of discipline-specific feedback observed in teacher feedback (Zhang et al., 2025). This ambivalence points to a larger struggle: students are willing to accept AI for grammar and mechanics but are wary of its ability to develop content and critical thinking.

Although these AI tools were faster and more consistent than humans in providing corrections, EFL students valued teacher feedback as being more reliable and contextual-

sensitive than AI feedback (Alnemrat et al., 2025). In a similar vein, Escalante et al. (2023) showed that although EFL learners appreciated AI-written feedback because it is transparent and accessible, they also feared that they are dependent on automated systems. This study is also evidence that human oversight and precision of AI-assisted feedback may be needed to instill a level of trust (Şimşek et al., 2025). Building on this, Mekheimer (2025) found that feedback using generative AI increased revision frequency and quality but needed to be critically filtered by learners. His study found that students trusted AI to identify surface-level issues with writing, but that trust eroded when suggestions contradicted what the student was trying to do with their writing (Mekheimer, 2025). Overall, these studies indicate that trust and perceived accuracy are contingent not only on the accuracy of the output but also on learners' critical-evaluation abilities and on the pedagogical environment in which they may learn from AI.

Perceived usefulness is one of the founding constructs of the Technology Acceptance Model, and newer studies show perceived usefulness to have remained salient surrounding AI feedback (Dias et al., 2022). In their meta-analysis of ChatGPT in educational contexts, Wang and Fan (2025) found that perceived usefulness was a stable predictor of learning performance and positive attitudes in different disciplines, including EFL writing. It went beyond technical functionality, such as feeling valued by learners, saving time, having the opportunity to iterate, and self-directed improvement (Wang & Fan, 2025). AI-generated feedback also led to statistically significant improvements in EFL learners' linguistic accuracy, which has an obvious outcome that demonstrates effective learning (Escalante, 2023). In the same way, Polakova and Ivenz (2024) found that students liked ChatGPT feedback in giving feedback on surface-level issues and also in higher-order aspects such as organization and coherence. They found that learners linked usefulness and the immediate and nonjudgmental nature of AI feedback, which helped give them autonomy during revision.

However, Werdiningsih et al. (2024) highlighted the irony of "perceived usefulness" and "genuine learner voice" and suggested that the use of "AI without questioning or resisting it" may be compromising original voice and ownership of creativity. Participants identified the utility of ChatGPT in the revision of drafts but worried that excessive use would take away from writing tasks that demand originality and critical engagement. This was in line with Alnemrat et al. (2025), who found that despite students feeling that AI feedback was effective, they still felt that teacher feedback provided more thorough contextual feedback. In this case, perceived usefulness is not only about getting it done, but AI feedback refines the humanity touch that exists in writing that feels like real-life academic writing.

The other consideration in implementing AI tools in EFL learning is the aspect of motivation. Yildiz (2023) states that ChatGPT reduced affective barriers, including writing anxiety, as it helped students with their motivation. The immediate and non-judgmental feedback provided by AI enhanced the desire of the learners to experiment in the field of writing. Similarly, Shao (2025) emphasized the significance of AI tools in creating a favorable emotional learning atmosphere by enhancing the level of self-efficacy and decreasing anxiety. These findings, as hypothesized in the Self-Determination Theory (Ryan and Deci, 2000), indicate a significant role played by the contextual support in helping intrinsic motivation. In particular, Shahat et al. (2025) showed that the mediator between the effect of using ChatGPT on the motivation of postgraduate students to conduct research and self-efficacy is the latter. Their study also

emphasized the beneficial effect that AI feedback may produce on the confidence students feel about their academic performance provided that the AI feedback may be perceived to be reliable, which leads to increased motivation to re-engage. Mekheimer (2025) additionally found that AI presented the aspect of revision through redoing that was a motivating factor and also resulted in the students becoming more engaged in the writing tasks.

Yet being motivated can be a two-sided sword. Similarly, Werdiningsih et al. (2024) found that while some students were motivated by ChatGPT feedback because the anxiety of feedback was relieved and AI generated feedback positively motivated students, helped them, and helped them in gaining confidence, others were demotivated and discouraged when they experienced the feedback as generic or poorly contextualized. This duality underlines the fact that the motivation is an AI brainchild giving feedback and the learner expecting the feedback. Overall, ChatGPT feedback can bring motivation to learn, reduce anxiety, and increase self-efficacy, although the effectiveness of the intervention will be determined by the possibility to meet the needs of the learners and the critical integration of the intervention with pedagogy.

Behavioral intention is the extent to which the learner will reuse ChatGPT to be provided with feedback. It is found that trust in combination with perceived usefulness and motivation determines such intention. Wang and Fan (2025) reported a strong association between perceived usefulness and behavior intention, affirming the predictive capacity of the Technology Acceptance Model. Similarly, Escalante (2023) noticed that students who achieved obvious linguistic performance improvements were more disposed to accepting AI feedback in future writing assignments. Polakova and Ivenz (2024) also found that sustained intention to use ChatGPT was highly influenced by how students perceived the tool as a reliable and supportive writing partner. In contrast, Alnemrat et al. (2025) found that doubts regarding feedback (accuracy and context sensitivity) decreased behavioral intention to act upon the feedback, but the effect was greater in students who had an overall greater trust in teacher guidance. According to Yıldız (2023), motivation also played an intervening role; students with higher motivation had greater intention to use AI tools for their editing. Overall, the findings suggest that the intentions for behavior are motivated by a combination of technology performance and affective and social-related constructs. Werdiningsih et al. (2024) found that students were reluctant to adopt ChatGPT due to moral issues in relation to authentication and plagiarism because of broader ethical conversations. Therefore, behavioral intention has to be understood in a broader context that includes not only trust, usefulness, and motivation, but also contextual and ethical boundaries.

The literature hints at the potential of ChatGPT feedback on the improvement of EFL academic writing. Research has suggested that students find the feedback beneficial and stimulating as well as develop better accuracy and a reduced need to revise the work frequently (Mekheimer, 2025; Polakova & Ivenz, 2024; Wang & Fan, 2025). At the same time, there are continuing challenges. Learners remain wary of accuracy and contextual depth as they continue to trust the feedback of their teachers more (Alnemrat et al., 2025; Escalante et al., 2023). Motivation is influenced both positively—through the reduction of anxiety and the improvement of confidence (Yildiz, 2023; Shao, 2025)—and negatively, as feedback is considered generic or inauthentic (Werdiningsih et al., 2024). Behavioral intention is ultimately dependent on the combination of the influence of trust, utility, and motivational outcomes.

Despite these advances, there are important gaps. Few studies have studied these four constructs (trust and accuracy, perceived usefulness, motivation and learning, and behavioral intention) together in a cohesive framework. Moreover, although the quantitative evidence is strong, qualitative aspects of learner reflections, such as reflecting on perceived strengths and suggestions for improving ChatGPT feedback, are understudied. Addressing these gaps using mixed-methods research will be helpful in understanding how ChatGPT can be integrated successfully as a sustainable feedback provider in CALL (computer-assisted language learning) contexts.

The current research is based on existing theories of technology acceptance, in particular the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT). TAM, which was proposed by Davis (1989), suggests that perceived usefulness and perceived ease of use influence the attitude toward technology and, in turn, behavioral intention. UNATAUT, developed by Venkatesh et al. (2003), further includes the other significant variables of social influence and facilitating conditions, and in the process, behavioral intention is mentioned as the most suitable predictor of actual use. These models offer a very relevant background for investigating how EFL students perceive and accept ChatGPT feedback (Venkatesh et al., 2003).

Within EFL writing, perceived usefulness is the degree of perceived usefulness or the perceived usefulness of using ChatGPT feedback to enhance their academic writing. Escalante (2023) demonstrated that artificial intelligence (AI)-generated feedback improved the accuracy of linguistic accuracy and revision practices of learners, which is a direct correspondence to the perceived usefulness construct of TAM. However, the addition of trust and perception of accuracy brings in a dimension that was not present in the original models. Youn et al. (2025) emphasized the notion that learners are more willing to adopt AI tools if they believe in the contextual accuracy and credibility of feedback, and thus trust can be seen as an extension of TAM and UTAUT in AI-mediated contexts.

Motivation and learning also offer significant additions to the framework by adding the affective and psychological dimensions that are not typically considered in models of acceptance. Zhai (2024) has demonstrated that grit, flow, and resilience moderate the relationship between the needs of learners and intention to use AI, which underscores motivation in the long-term engagement. Motivation in writing enhances not only the early adoption of AI, but also other revision practices (Zhai et al., 2024). Alwasidi and Al-Khalifah (2025) also affirmed that ChatGPT feedback reinforced the confidence and perseverance of the learners, which, in turn, enhanced the desire to use AI assistance in writing.

When consolidated, this framework proposes that the adoption of ChatGPT takes place in a dynamic manner that is influenced by perceived usefulness, trust and accuracy, and motivation, each being related to predicting behavioral intention. Perceived usefulness represents identification of the usefulness of this tool on the improvement of writing performance, whereas trust represents the credibility and reliability of the feedback received from this tool. Motivation can be conceptualized as the cognitive and emotional forces that maintain involvement with AI-assisted revision. In doing so, the model combines issues by extending the limits of TAM and UTAUT as the structure informing our study to elaborate how students assess the process and acceptance of ChatGPT as a feedback provider.

In this study we aim to explore EFL students' perceptions toward the use of ChatGPT as a feedback giver for academic writing. We specifically investigate the evaluative usability, credibility, quality, and motivational transference of ChatGPT-generated feedback, as well as the corresponding effect on students' intention to continue using the tool. Second, we explore the qualitative reflections of students in going beyond their initial expectations and experiences. We sequentially integrate quantitative and qualitative methods to enrich our understanding of the role of ChatGPT in supporting EFL writing. We also enhance the Unified Theory of Acceptance and Use of Technology (UTAUT)-based theoretical models by extending theoretical models of technology acceptance to the context of AI-mediated feedback.

METHOD

For investigating EFL learners' experiences with ChatGPT-generated feedback, we incorporated a mixed-method design by integrating both qualitative and quantitative methods. We opted for a mixed-method approach to understand the entirety of the phenomenon as it integrates strengths of both qualitative and quantitative methods. Moreover, it also aids in exploring both the depth and breadth of the phenomenon. This improves the validity and richness of my interpretation of the study. We used an explanatory sequential model, in which we collected quantitative data and analyzed them before adding the qualitative component, with which we aimed to extend and contextualize the findings (Fetters et al., 2013).

Participants included 254 undergraduate EFL students enrolled in academic English writing courses at a Middle Eastern university. The sample was made up of males and females of diverse disciplines with proficiency levels from intermediate to advance. Since students had previously been exposed to AI assisted writing tools like ChatGPT, we used a convenience sampling strategy. Participation was voluntary, and all students provided their informed consent. To ensure diversity of subjects, a few of the participants had never used ChatGPT before (Dornyei, 2007).

Quantitative Survey: The major tool was the questionnaire in the form of 24 item scales that were adapted from previous well-established models, namely, TAM (Davis, 1989) and UTAUT (Venkatesh et al., 2003). Four constructs were measured: usefulness perception, trust and accuracy, motivation and learning, and behavioral intention. A 5-point Likert—1 strongly disagree and 5 strongly agree—scale was applied to the items. The run did the reliability test, and the results were a .83 (cronbach's alpha), which shows high internal consistency (Pallant, 2020). *Qualitative Open-Ended Questions:* To obtain more in-depth insights into student experiences, we included three open-ended items: (a) student surprises from ChatGPT feedback, (b) ChatGPT feedback style impressions, and (c) student improvement suggestions. The goal of these questions was to probe ChatGPT's strength and limitations as well as the pedagogical role.

In the spring semester of 2025, we got the chance to gather data on a weekly basis and took a span of six weeks. During the first phase, the questionnaire was completed after the students completed writing tasks and obtained ChatGPT feedback. During phase two, 75 respondents answered open-ended questions in a bid to triangulate the quantitative and qualitative conclusions. This design enhanced both validity and reliability (Plano Clark, 2017).

The software, SPSS helped us to analyze survey data. To conclude the student perceptions on four constructs, we initially calculated descriptive statistics for each.

Cronbach's alpha was once again tested with correlation and regression to test the relationship between the variables. The hypothesis testing involved mediation tests with Process Macro developed by Hayes (2022) (Model 4) to determine whether motivation mediated the effect of perception of usefulness of behavioral intention. Effect sizes were also computed to give strong evidence of practical significance (Hayes, 2017).

Qualitative Analysis: The responses were coded by a thematic analysis (Braun and Clarke, 2006) according to a six-phase process (Braun and Clarke, 2006). We came up with codes and categories, and as we put together categories, major themes came out.

In the course of conducting the research, we incorporated ethical considerations provided by the British Educational Research Association (BERA, 2018). Prior to data collection, the university Institutional Review Board ethical approval was obtained. Students were guaranteed confidentiality and anonymity and assured their participation would not impact their academic records.

In our data analysis, we provide impressions, both qualitative and quantitative. We have divided this section into two sections, the statistical patterns through SPSS and the perceptions of learners through thematic analysis. This twofold strategy will guarantee the deeper insight and verification of the results by the means of both numeric data and open-ended responses of the participants.

RESULTS AND DISCUSSION

Descriptive statistics of demographic variables; there were 254 students involved in the study, as in Table 1. The sample was skewed towards males who constituted a little more than 71 percent of the sample, whereas females constituted 28.7 percent. Over half of the sample (53.5%) were between the ages of 20-29, with 24.4% aged 30-39, 18.1% younger than 20, and 3.9% aged 40-49, respectively. Undergraduates took the highest proportion of 44.9%, followed by graduates (31.9%) and postgraduates (23.2%) in that order. English proficiency was also found to be mostly high (89.8), with just 10.2 percent reporting to be intermediate. Almost everyone (99.6%) expressed having used ChatGPT to write comments in the past, which proved that the sample was quite conversant with the application.

Reliability statistics: the Cronbach's alpha of the total 24-item scale was .83 (excellent range), indicating that the instrument has excellent internal consistency. It validates the idea that the modified tool was suitable in gauging the perceptions of students about the ChatGPT feedback in the present sample.

Table 1. Demographic Characteristics of Respondents

Variable	Category	SD	%
Gender	Male	181	71.3
	Female	73	28.7
Age	Under 20	46	18.1
	20-29	136	53.5
	30-39	62	24.4
	40-49	10	3.9
Level of study	Undergraduate	114	44.9
	Graduate	81	31.9
	Post Graduate	114	23.2
English Proficiency	Intermediate	26	10.2
	Advanced	228	89.8
ChatGPT usage	Yes	253	99.6
	No	1	0.4

Perceptions of the Usefulness of ChatGPT Feedback

To summarize the responses of students on the usefulness scale, we applied objective descriptive statistics, and to compare the mean score of usefulness, we applied a one-sample t-test. We used the d value of Cohen to estimate the size of the difference.

Table 2. Descriptive Statistics for Perceived Usefulness Items

Item	Mean	SD
Grammar improvement	4.38	0.51
Organization of idea	4.37	0.53
Error explanation clarity	4.35	0.52
Overall improvement	4.25	0.53
Faster revision	4.37	0.53
Alignment with academic standards	4.35	0.52
Overall PU	4.34	0.51

Table 2 of the descriptive statistics indicates that students rated ChatGPT feedback as useful reliably, with the items all showing a mean greater than 4.20. Grammar improvement and organization of ideas had the highest rating, which means that students rated ChatGPT especially high on assisting with sentence correction and general text flow.

Table 3. One-Sample T-Test for Overall Perceived Usefulness (PU)

Variable	Mean	SD	t (df)	p	Choen's d
PU	4.34	0.31	223.69(253)	<0.001	14.04

A one-sample t-test established that the mean score was much higher than the neutral middle of 3, $t(253) = 223.69$, $p < .001$ as indicated in Table 3. This finding demonstrates that the learners were not only neutral but also were immensely helpful in receiving ChatGPT feedback and using it in academic writing. The effect size for the perceived usefulness scale was extremely large (Cohen's $d = 14.04$), indicating that the difference between neutrality and students' positive perceptions was not only statistically significant but also of major practical importance. Finally, H1 is supported, since students' ratings were significantly above the neutral point, with a very large effect size demonstrating strong practical significance.

Trust and Accuracy of ChatGPT Feedback

We incorporated descriptive statistics, a one-sample t-test, and Cohen's d to investigate and justify the second objective

Table 4. Descriptive Statistics for Trust and Accuracy (TA) Items

Item	Mean	SD
Accuracy of feedback	4.28	0.53
Context understanding	4.35	0.56
Confidence in using feedback	4.40	0.52
Comparison with teacher feedback	4.40	0.52
Task-dependent performance	4.35	0.56
Trust in informal vs formal writing	4.40	0.52
Overall TA	4.36	0.31

Descriptive analysis revealed that students rated ChatGPT feedback as highly trustworthy, with all items' means above 4.20 (as per Table 4). The strongest agreement was for confidence in using ChatGPT feedback and trust when compared with teacher feedback, suggesting that students generally relied on ChatGPT for writing guidance.

Table 5. One-Sample T-Test for Overall Trust and Accuracy

Variable	Mean	SD	t (df)	p	Choen's d
TA	4.36	0.31	222.41(253)	<0.001	13.96

As depicted in Table 6, the one-sample t-test showed that the mean trust score was significantly higher than the neutral midpoint of 3, $t(253) = 222.41$, $p < .001$. This affirms that students rated the feedback given by ChatGPT as much more reliable and correct as compared to a neutral rating. The effect size was so large (it is depicted in Table 5), and student trust in ChatGPT feedback was not only statistically significant but also practically meaningful. Therefore, H_2 is accepted since the t-test statistics and extremely high effect size are good evidence that students strongly believe in the accuracy and relevance of writing feedback by ChatGPT.

Predictors of Behavioral Intention

To analyze the third objective, we used Pearson and multiple regression tests to evaluate the overall predictive value of usefulness, trust, and motivation on behavioral intention. Lastly, we tested mediation with Hayes's PROCESS macro, in which we tested whether motivation mediated the relationship between intention and usefulness.

Correlation Analysis

Table 6. Correlations among Key Study Variables

Variable	PU	TA	ML	BI
PU	1	.385**	.228**	.750**
TA	.385**	1	.425**	.587**
ML	.228**	.425**	1	.651**
BI	.750**	.587**	.651**	1

Table 6's correlation analysis showed that students who rated ChatGPT feedback more useful, trustworthy, and motivating also reported a greater desire to continue using it as a provider of feedback. Perceived usefulness and behavioral intention had the best correlation, which suggests that the perception of the practical utility of ChatGPT feedback is closely related to the desire to use it in the future.

Regression Analysis

Table 7. Model Summary for Predictors of Behavioral Intention

R	R ²	Adjusted R ²	Std. Error of Estimate
0.910	0.828	.826	.115

The regression model explained 82.8% of the variation in behavioral intention, indicating that the overall perceptions of ChatGPT feedback (usefulness, trust, and motivation) are strong predictors of whether students will go on to use it as a feedback tool in the future, as shown in Table 7.

Table 8. Regression of Coefficients for Predictors of Behavioral Intention

Predictor	B	SE	B (Standardized)	T	P
PU	0.519	0.025	.582	20.45	<.001
TA	0.154	0.027	.174	5.69	<.001
ML	.373	0.024	.444	15.30	<.001

The three predictors were all statistical in Table 8. The greatest motivator of intention was use, which indicates that students will most likely keep using ChatGPT when they find the feedback highly useful in enhancing their writing. Motivation was also a significant factor, as it demonstrated that the ChatGPT feedback can encourage students to remain active in revision. Although it was less prominent, trust still affected intention to a considerable degree, which shows that to continue using ChatGPT, one should trust that its feedback is accurate as well.

Further, Figure 1 presents the unique contribution of the perceived usefulness to behavioral intention when other predictors have been controlled in the model. The slope of the regression line is positive and significant, and R^2 indicates that a large percentage of the variance in behavioral intention is predicted by usefulness. The story supports that students who find ChatGPT feedback more valuable tend to switch to using it significantly more often, and perceptions of usefulness are the best predictor of the studied variables.

Figure 1. Partial Regression Plot of Perceived Usefulness Predicting Behavioral Intention

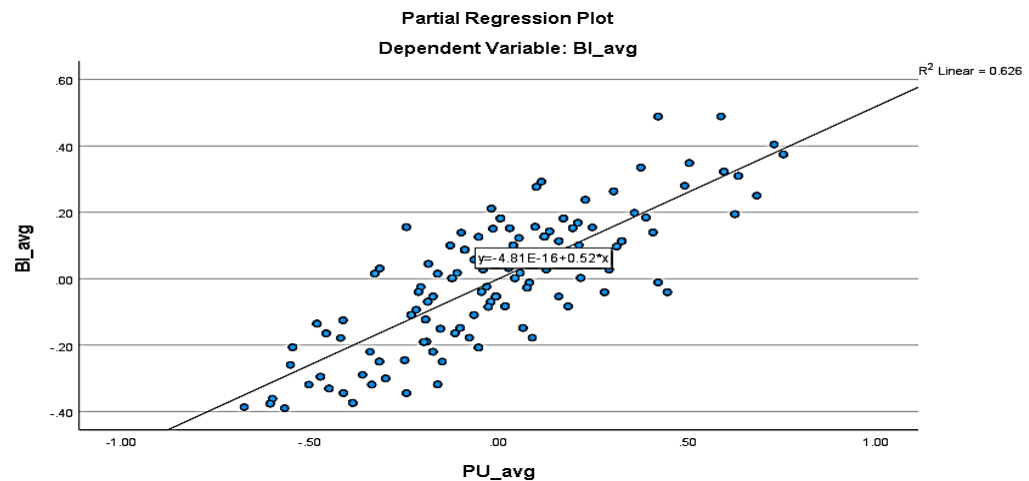


Figure 2 shows the only influence of trust and accuracy on behavioral intention when the impact of other predictors has been removed in the model. The positive slope shows that the increased the confidence in the feedback provided by ChatGPT, the more the intention to use it, which demonstrates that the confidence in the quality of AI feedback has a significant impact on the desire to use it in writing school papers. Figure 3 shows the special role of motivation in behavioral intention, and adjusts other predictors in the regression equation. The slope is positive, and the value of R^2 is good evidence of a strong positive association since higher motivation levels significantly contribute to an intention to use ChatGPT feedback. The predictive power of motivation has been proven by the plot, which underpins its definition as a major independent predictor and a mediator in the context of the research.

Figure 2. Partial Regression Plot of Trust and Accuracy Predicting Behavioral Intention

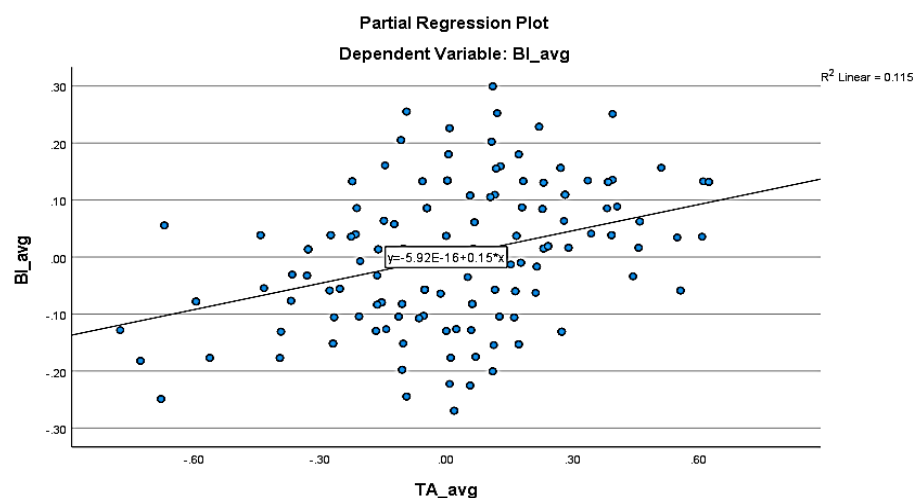
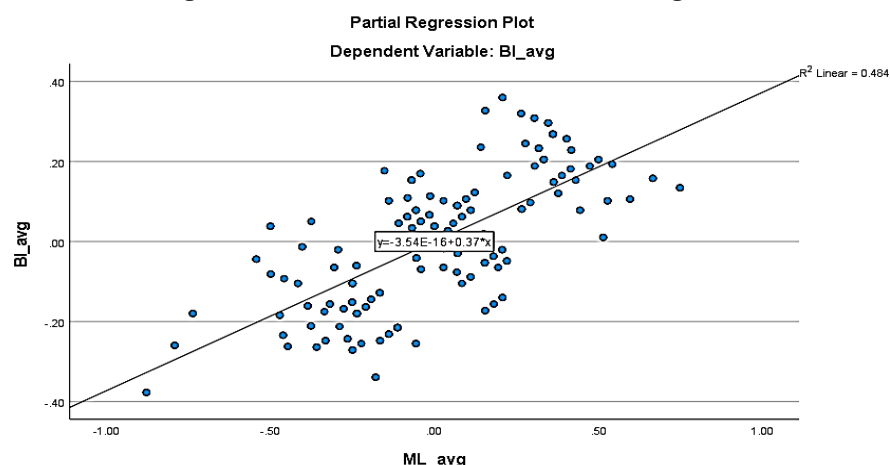
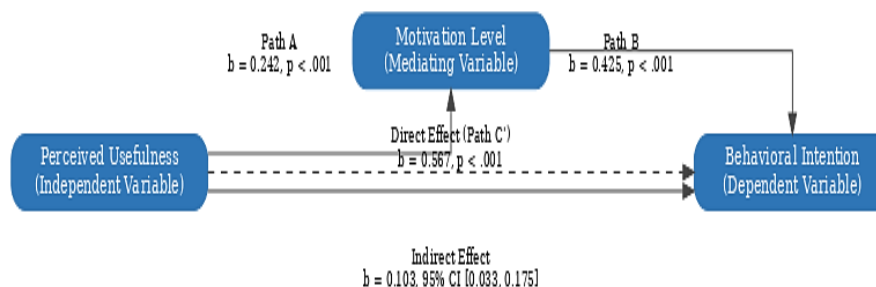


Figure 3. Partial Regression Plot of Motivation Level Predicting Behavioral Intention**Table 9. Direct and Indirect Effects of Perceived Usefulness on Behavioral Intention**

Effect	B	SE	95% CI [LL, UL]	SIG
Direct Effect (PU→BI)	0.567	0.026	[0.516, 0.617]	p<.001
Indirect Effect (PU→ML→BI)	0.103	0.037	[0.033, 0.175]	CI Excludes 0

Mediation analysis showed that perceived usefulness had both direct and indirect effects on behavioral intention via motivation. The direct effect ($B = 0.567$, $p < .001$) indicates the fact that students who believe that the feedback of ChatGPT is helpful are more likely to use it again. The indirect effect ($B = 0.103$, 95% CI [0.033, 0.175]) shows that usefulness also increases the motivation of students, and it consequently increases their intention to use ChatGPT feedback. This validates a partial mediation effect where usefulness does not only make the students believe in the value of ChatGPT, but it also motivates them to remain active in its feedback. These results demonstrate a strong reason to believe in H3: the behavior intention of the students to use ChatGPT in the future basically depends on how useful, trustworthy, and motivating they believe the feedback of the program to be.

Furthermore, as Figure 4 demonstrates, the perceived usefulness is the strongest predictor of motivation (Path A, $b = 0.242$, $p < .001$) and behavioral intention both directly (Path C', $b = 0.567$, $p < .001$) and indirectly through motivation (Path B, $b = 0.425$, $p < .001$; indirect effect $b = 0.103$, 95% CI = [0.033, 0.175]). This implies partial mediation, that is, motivation enhances the association between helpful perceptions and intention to remain using ChatGPT feedback.

Figure 4. Mediation Model of Motivation Between Perceived Usefulness and Behavioral Intention

To address this final study objective, we thematically analyzed the open-ended responses of students. We coded and placed patterns in categories and generalized into the broader themes on each of the three open-ended questions.

Surprising Insights from Using ChatGPT Feedback

Three themes in responses were deduced, as discussed below. These themes are depicted in Figure 5.

Theme 1: Heightened Awareness of Language Errors

Most students were surprised by the effectiveness of ChatGPT to identify minor yet quantifiable grammatical or lexical errors that they have frequently missed. This caused more awareness of the patterns in their writing and promoted tightening of their self-monitoring.

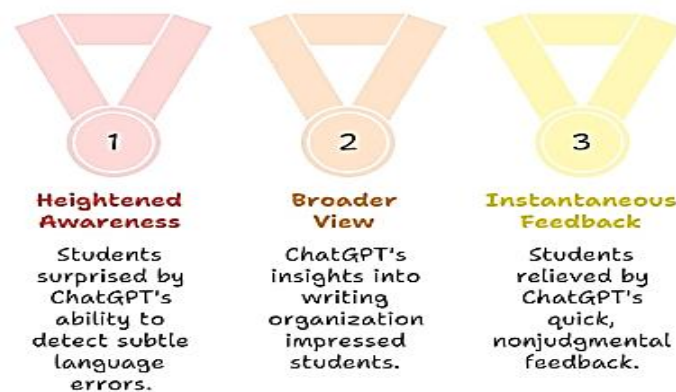
Theme 2: Broader View of Writing Organization

The other surprise was that ChatGPT indicated the flow, transition, and paragraph coherence problems. Students had anticipated grammar level corrections but were astonished by its ability to attract focus to the general structure and clarity.

Theme 3: Instantaneous, Nonjudgmental Feedback

Some of the reflections focused on the surprise of getting feedback with a non-critical attitude. According to students, ChatGPT was able to provide fast and detailed recommendations in a non-authoritative manner, which allowed the revision process to be less stressful than a regular feedback experience.

Figure 5. Surprising Insights from Using ChatGPT Feedback
Top Surprises in ChatGPT Feedback



Personality of ChatGPT's Feedback Style

Three distinct themes emerged from the responses, discussed in detail below. These themes are visually represented in Figure 6.

Theme 1: A Friendly Tutor

Most often, it was characterized as a friendly tutor, ChatGPT. Students emphasized in the positive tone, clear descriptions, and non-scary instructions that they were prompted to work on revisions.

Theme 2: A Helpful Peer

Other students likened it to a friendly peer reviewer, and even though ChatGPT was not as authoritative as a teacher, it provided useful and supportive feedback that was collaborative, but not prescriptive.

Theme 3: A Balanced Teacher-Like Role

A fewer number called ChatGPT a teacher who is a balanced teacher, firm and yet not too strict at the same time. Students claimed that this balance had led to trust in the dependability of the feedback.

Figure 6. Personality of ChatGPT's Feedback Style
How does ChatGPT's feedback style resonate with you?



Suggested Changes for ChatGPT Feedback

The answers to the last questions provided resulted in the three key themes identified as described below. Figure 7 shows these themes.

Theme 1: Greater Contextual Sensitivity

Students suggested that ChatGPT should be more specific to the academic situation, i.e., disciplinary conventions or the requirements set by the assignment, instead of using a general writing standard.

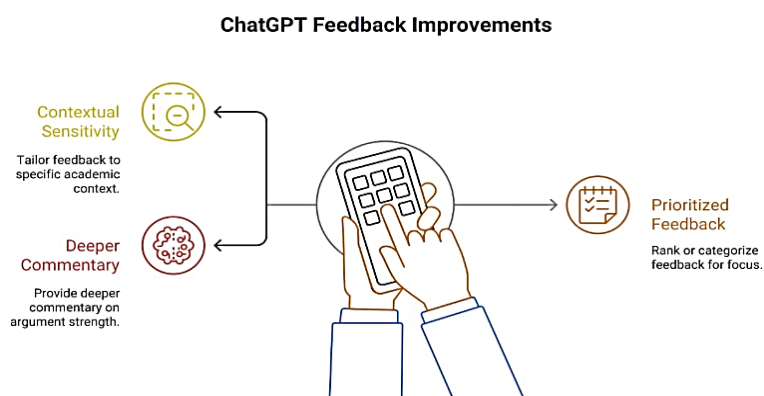
Theme 2: Clearer Prioritization of Feedback

A number of the reflections were called ranked or categorized feedback (e.g., grammar vs. organization vs. content relevance). Students thought that prioritization could guide them to first concentrate on the most important issues.

Theme 3: Expansion Beyond Surface-Level Corrections

Participants also wanted ChatGPT to be more communicative, particularly with regard to argument strength, critical engagement, or theoretical framing, which it still does not do as well as human instructors.

Figure 7. Suggested Changes for ChatGPT Feedback



The unexpected revelations indicate how ChatGPT redefined the expectations of learners of feedback in its ability to not only respond to grammar but also create awareness of form and offer a nonjudgmental response. The personality feedback shows that students found ChatGPT helpful, whether it is a friendly tutor or a friend, which is corroborated by quantitative data that motivation is a robust predictor of behavioral intention. Lastly, the proposed solutions correspond to the issues of trust and accuracy since students indicated that they prefer more contextual, prioritized, and sophisticated feedback. Students noted the usefulness of ChatGPT as a feedback-giving tool but noted some limitations to its feedback that still exists relative to that offered by human teachers.

The direct quotes of the respondents, sorted into three main themes, can be seen in Figure 8.

Figure 8. Themes from Students' Qualitative Reflections on ChatGPT Feedback

Q1: SURPRISING INSIGHTS FROM USING CHATGPT FEEDBACK	Q2: PERSONALITY OF CHATGPT'S FEEDBACK STYLE	Q3: SUGGESTED CHANGES FOR CHATGPT FEEDBACK
i). I was surprised at how often I repeated the same grammar mistake, and ChatGPT pointed it out every time. ii). It made me realize I use too many passive sentences, something I had never noticed before. iii). I expected grammar corrections, but it actually helped me organize my paragraphs more clearly. iv). ChatGPT highlighted weak transitions, which no teacher had ever pointed out to me. v). It gives feedback instantly, without the stress of waiting for my teacher's comments. vi). The feedback was supportive and never judgmental, which encouraged me to revise.	i). It feels like a friendly tutor guiding me step by step. ii). The tone is encouraging, like someone is cheering me to keep improving. iii). Sometimes it feels like working with a peer who gives suggestions rather than orders. iv). It is like a classmate reviewing my work—casual but still useful. v). It points out errors firmly but not in a harsh way, more like a patient teacher. vi). The feedback has authority, yet it is not strict, which makes me trust it more.	i). It should give feedback more suited to my assignment topic, not just general rules. ii). Sometimes it misunderstands the task—more context awareness would help. iii). I want it to separate grammar mistakes from organization advice so I know where to start. iv). It should rank errors by importance, like first fix structure, then grammar. v). ChatGPT is great at grammar, but I wish it commented on my argument strength too. vi). It rarely tells me if my ideas are logical; it only checks language.

Triangulation of Quantitative and Qualitative Findings

The combination of quantitative and qualitative results provides the complete description of the way students engage with the ChatGPT feedback. The quantitative outcome validated the high levels of perceived usefulness (RO1), moderate trust in accuracy (RO2), and good correlations between usefulness, motivation, and behavioral intention (RO3). Qualitative reflections on these outcomes were that ChatGPT is effective in grammar correction and text organization, but that there were issues with contextual accuracy and level of criticism. Some described the tool as a peer or tutor presence, which matches its motivational purpose revealed in regression analysis. Meanwhile, appeals to more context-sensitive and prioritized feedback indicate constraints of trust, which became a less significant predictor of intention. These two data strands indicate that ChatGPT is the additional feedback on top of the human feedback as opposed to a substitute.

CONCLUSION

Students' perspectives about AI-generated feedback were measured across four dimensions, i.e., usefulness, trust and accuracy, motivation, and behavioral intentions. Students perceived ChatGPT as a nonjudgmental partner in learning, as it provides immediate, clear, and definite responses that improve linguistic accuracy and reduce the anxiety. Learners have not only praised the supportive role of ChatGPT but also raised questions about its authenticity, contextual accuracy, ethical usage, and occasional inability to process argument or discipline-specific conventions. From our findings, it can be inferred that AI-generated feedback must be accompanied and critically gauged by human oversight. The amalgamation of qualitative and quantitative findings indicates that ChatGPT cannot replace traditional feedback in its entirety, though it can complement the human feedback.

We divided the analysis contents into four groups, including the usefulness of ChatGPT feedback, trust, motivation, and behavior intention among EFL students. The

findings were as follows: the learners rated it as immediate, direct, and non-judgmental in response, which resulted in more linguistic accuracy, less anxiety, and more revisions and supported its activity as a non-judgmental writing collaborator. However, students questioned the contextual validity, authenticity, and ethical use of measures of higher-order processes that include argumentation and disciplinary conventions. It implies that AI must be treated with seriousness when giving feedback, and it must be complemented with the mentoring of humans as the findings describe. Combining both quantitative and qualitative data, the results indicate that ChatGPT is a complement rather than a substitute for traditional feedback, and it is possible to extend the existing theories on technology acceptance that trust and motivation are determinant factors influencing the adoption of technology.

Despite the contributions, this research is limited in one way or another. The greatest weakness is that there is no generalizability because the sample was picked in a single university, and therefore the results cannot be extended to different universities and cultures. The other weakness is bias since the learners will be self-reporting, and this exposes them to bias. Additionally, an immediate impact on the perception related to the close experience with ChatGPT may be short-term because no study was performed regarding the impact of the experience on writing development and the long-term intention to behave this way.

The longitudinal designs exceeding one semester may also provide insight into how the feedback may be designed, taking into consideration the dependence on the ChatGPT feedback with time, and thus, provide the indication of the possible future direction of the research. More generalizability is produced due to comparative studies done on different institutions and levels of proficiency, as can be gained when the experimental design is applied to control groups of just the teacher feedback. Moreover, the analysis of the context and writing that is genre-specific can also reveal whether the ChatGPT replies are as helpful outside of the wide scope of academic writing. These contexts will be instrumental in the comprehension of how the relevance of ChatGPT is changing, let alone how this technology can be implemented in a responsible way in EFL writing pedagogy.

REFERENCES

- Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: a quantitative study. *Frontiers in Education*, 10, 1614673. <https://doi.org/10.3389/feduc.2025.1614673>
- Alwasidi, M. A., & Al-Khalifah, K. S. (2025). Assessing the impact of ChatGPT on EFL students' writing productivity and proficiency. *Journal of Language Teaching and Research*, 16(3), 986–995. <https://doi.org/10.17507/jltr.1603.18>
- Bera (2018). Ethical guidelines for educational research. Retrieved January 10, 2026, from https://study.sagepub.com/sites/default/files/bera_ethical_guidelines_2018_4th_ed.pdf
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>

- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Dias, S. B., Oikonomidis, Y., Diniz, J. A., Baptista, F., Carnide, F., Bensenousi, A., Botana, J. M., Tsatsou, D., Stefanidis, K., & Gymnopoulos, L. (2022). Users' perspective on the AI-Based smartphone PROTEIN App for personalized nutrition and healthy living: a modified technology acceptance model (mTAM) approach. *Frontiers in Nutrition*, 9, 898031.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1), 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—Principles and practices. *Health Services Research*, 48(6 Pt 2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.
- Kurt, G., & Kurt, Y. (2024). Enhancing L2 writing skills: ChatGPT as an automated feedback tool. *Journal of Information Technology Education: Research*, 23, 1–17. <https://doi.org/10.28945/5370>
- Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4(1), 170. <https://doi.org/10.1007/s44217-025-00602-7>
- Mizumoto, A., Shintani, N., Sasaki, M., & Teng, M. F. (2024). Testing the viability of ChatGPT as a companion in L2 writing accuracy assessment. *Research Methods in Applied Linguistics*, 3(2), 100116. <https://doi.org/10.1016/j.rmal.2024.100116>
- Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' perceptions of generative artificial intelligence (GenAI) use in academic writing in English as a foreign language. *Education Sciences*, 15(5), 611. <https://doi.org/10.3390/educsci15050611>
- Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS*. Routledge. <https://doi.org/10.4324/9781003117452>
- Plano Clark, V. L. (2017). Mixed methods research. *The Journal of Positive Psychology*, 12(3), 305–306. <https://doi.org/10.1080/17439760.2016.1262612>
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Shahat, H. A., Badawy, M. E., Elballah, K. A., & Ibrahim-Shook, A. (2025). Self-efficacy as a mediator between ChatGPT usage and research motivation among postgraduate students. *International Journal of Innovative Research and Scientific Studies*, 8(4), 987–996. <https://doi.org/10.53894/ijirss.v8i4.7982>

- Shao, S. (2025). The role of AI tools on EFL students' motivation, self-efficacy, and anxiety: Through the lens of control-value theory. *Learning and Motivation*, 91, 102154. <https://doi.org/10.1016/j.lmot.2025.102154>
- Şimşek, A. S., Cengiz, G. Ş. T., & Bal, M. (2025). Extending the TAM framework: Exploring learning motivation and agility in educational adoption of generative AI. *Education and Information Technologies*, in press. <https://doi.org/10.1007/s10639-025-13591-9>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*, 12(1), 1–21. <https://doi.org/10.1057/s41599-025-04787-y>
- Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities*, 11(1), 2392388. <https://doi.org/10.1080/23311983.2024.2392388>
- Yıldız, T. A. (2023). The impact of ChatGPT on language learners' motivation. *Journal of Teacher Education and Lifelong Learning*, 5(2), 582–597.
- Youn, C. H., Salam, A. R., & Rahman, A. A. (2025). AI-driven tools in providing feedback on students' writing. *International Journal of Research and Innovation in Social Science*, 9(3s), 58–67.
- Zhai, X., Zhao, R., Jiang, Y., & Wu, H. (2024). Unpacking the dynamics of AI-based language learning: Flow, grit, and resilience in Chinese EFL contexts. *Behavioral Sciences*, 14(9), 838. <https://doi.org/10.3390/bs14090838>
- Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T., & Cukurova, M. (2025). Evaluating trust in AI, human, and co-produced feedback among undergraduate students. *arXiv preprint*, arXiv:2504.10961. <https://doi.org/10.48550/arXiv.2504.10961>