SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 6431-6442
Open Access. Online & Print

DOI: 10.5281/zenodo.19511702

ARTIFICIAL INTELLIGENCE AND LEARNER ENGAGEMENT IN SAUDI EFL WRITING CLASSROOMS: A LONGITUDINAL LATENT GROWTH CURVE STUDY

Amera Alharbi[1] and Turki Alharbi[2]

[1]Saudi Electronic University Department of English Language and Translation. Email: aa.alharbi@seu.edu
Orcid ID: https://orcid.org/0000-0002-3799-8543
[2]Saudi Electronic University Department of English Language and Translation. Email: t.alharbi@seu.edu.sa
Orcid ID: https://orcid.org/0009-0005-1874-1557

Received: 21/01/2026
Accepted: 25/03/2026

Corresponding Author: Amera Alharbi (aa.alharbi@seu.edu.sa)

ABSTRACT

The rapid integration of AI tools in higher education is reshaping writing instruction, yet their longitudinal impact on learner engagement remains insufficiently understood. This study tracked 55 first-year Saudi undergraduate EFL students across a six-week AI-assisted writing course, collecting data at three points: before the intervention, at the midpoint, and upon completion. To examine how engagement evolved across cognitive, behavioral, emotional, and agentic dimensions, Latent Growth Curve Modelling (LGCM) was employed using AMOS 31. All four dimensions of engagement increased significantly, with cognitive and agentic engagement showing the steepest trajectories. Technological proficiency emerged as a consistent differentiator: students with stronger digital skills engaged with AI tools more purposefully, revising thoughtfully and self-regulating more effectively. The benefits of AI-supported writing environments depend substantially on the digital competence students bring to their learning. Building technological readiness alongside AI integration is not optional — it is foundational to equitable learning outcomes in the digital university.

KEYWORDS: Learner Engagement, AI-Assisted Writing, Latent Growth Curve Modelling, Technological Proficiency, EFL Writing, Saudi Higher Education.

Copyright: © 2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License. (https://cre-ativecommons.org/licenses/by/4.0/).

Restricted - مقيد

1. INTRODUCTION

The rapid incorporation of AI writing tools has opened new opportunities and new uncertainties in equal measure. These tools have moved from supplementary aids to integral classroom resources, raising a fundamental pedagogical question: do they produce lasting growth across the multiple dimensions of learner engagement, or merely a short-term novelty effect? The answer has direct implications for how institutions should invest in AI integration — as a one-time enhancement or as a sustained instructional commitment requiring structural support.

Engagement, defined by what learners do, how they feel, how they think, and how they take charge of their learning, has been found to be one of the strongest predictors of learners' academic success and writing development (Reeve & Tseng, 2011; Fredricks et al., 2004). This challenge is particularly acute for EFL learners, whose writing-related anxiety and cognitive load can suppress engagement even among those who are intrinsically motivated (Haji, 2022; Li & Fan, 2024). Researchers have argued that AI tools support engagement by providing immediate, non-judgmental feedback that allows learners to experiment with language without fear of criticism (Luo & Yusuf, 2025; Yan & Zhang, 2024). Theoretically, the argument is well-founded. What has been lacking is rigorous longitudinal evidence.

The majority of studies have employed a snapshot approach to examining the effects of these tools on learners' engagement. Such designs cannot determine whether engagement was sustained over time, or whether learners developed genuine autonomy or passive dependency. Nor can it tell us how individual differences, particularly learners' digital readiness, would affect these trends of engagement (Liu et al., 2025; Zhou et al., 2023). The current study directly addresses this gap. By applying a three-wave longitudinal design and using LGCM, we investigated the development of learners' cognitive, behavioral, emotional, and agentic engagement over a six-week AI-assisted writing course among 55 first-year Saudi undergraduates and how learners' technological proficiency explained individual variation. Three research questions guided the study:

RQ1: Do learner engagement dimensions change significantly across three time points following systematic AI writing tool use?

RQ2: What are the specific latent growth trajectories (linear or non-linear) of these engagement dimensions over the course of an academic semester?

RQ3: To what extent does prior technological proficiency explain individual differences in the initial levels (intercepts) and the rates of growth (slopes) of learner engagement? To situate these questions theoretically, the following section reviews the conceptual frameworks underpinning the study's design and analytical approach.

2. THEORETICAL FRAMEWORK

2.1. The Four-Dimensional Engagement Framework

This study adopts Reeve and Tseng's (2011) four-dimensional model, conceptualising learner engagement as comprising behavioral (effort, persistence, participation), emotional (motivation, confidence, affective investment), cognitive (strategic thinking, deep processing, metacognitive monitoring), and agentic (proactive goal-setting, self-direction) components. This framework is well suited to AI-mediated writing because it captures the full spectrum of student response — from surface compliance to strategic self-regulation and genuine ownership of the writing process. Reeve et al.'s (2020) later extension of the model confirmed that engagement is a fluid state that is responsive to contextual factors, thus making it a state that must be longitudinally measured.

2.2. Longitudinal Perspectives And LGCM

Studies within the technology-enhanced learning environment have uniformly confirmed that student engagement shifts dynamically as learners grow more familiar with the technology and the feedback they receive (Bergdahl et al., 2020; Henrie et al., 2015). While initial enthusiasm for new technology may plateau for some students, others follow a trajectory from low to high engagement as they build competence. LGCM mirrors this developmental perspective: it estimates both initial status (intercept) and rate of change (slope), and crucially reveals individual variation in those parameters — making it particularly suited to revealing not just whether students improved, but who improved most and why (Grimm & Ram, 2018; Zhang, 2022).

2.3. Technological Proficiency as a Moderator

Technological proficiency—the capability to apply, interpret, and effectively employ digital tools for educational purposes—has become a noteworthy difference variable in technology-enhanced language learning (Zhou & Wei, 2018). In AI-mediated writing, its role is enhanced because students must not only understand AI feedback linguistically but also assess its relevance, determine what to incorporate, and

Restricted - مقيد

synthesize suggestions into a cohesive revision strategy. Students who do not possess these skills may become passive consumers of automated output instead of active learners (Xie et al., 2023; Shen et al., 2023). Sociocultural perspectives position technology as a mediating instrument for intentional learning (Vygotsky, 1978; Lantolf & Thorne, 2006), whereas self-determination theory associates perceived competence with intrinsic motivation and ongoing engagement (Reeve et al., 2020). Together, these frameworks predict an amplification effect: technologically proficient learners will not only begin at higher engagement levels but grow faster — a hypothesis this study tests directly. These theoretical predictions form the basis of the analytical framework applied in the methodology described below.

3. METHODOLOGY

3.1. Research Design

The current study used a quantitative longitudinal design with three-time stages in a six-week AI-assisted writing course. The choice of a longitudinal design was deliberate and theoretically motivated. Engagement is a dynamic process that unfolds in response to instructional experiences, not a static quality that can be measured at a single point in time (Reeve et al., 2020; Wong & Liem, 2021). Cross-sectional designs, which dominate the existing literature on AI and EFL writing, are structurally incapable of answering whether engagement gains persist, accumulate, or fade over the course of instruction. A three-wave longitudinal design allows precisely that — it maps a developmental trajectory rather than a static position.

Three time points were chosen to balance methodological rigour with practical feasibility. Two measurement waves would permit only a pre-post comparison, making it impossible to distinguish a linear growth pattern from a nonlinear one (e.g., rapid early gains followed by a plateau). Three waves provide the minimum number of data points required to fit a latent growth model with identified intercept and slope parameters, while remaining achievable within a single-semester instructional period without excessive participant burden. The six-week window was determined by the length of the writing courses in which participants were enrolled, making the design ecologically valid — data collection was embedded within real instructional conditions rather than imposed on a separate experimental protocol.

Two complementary analytical approaches were used in tandem. Repeated-measures ANOVA, corrected for the confirmed sphericity violation using the Greenhouse-Geisser adjustment, examined overall group-level change over time and identified which specific intervals produced significant differences. LGCM then went further — modelling not just average trajectories but individual variation around those trajectories, estimating each participant's latent intercept (initial status) and latent slope (rate of change), and testing whether prior technological proficiency predicted those individual parameters. This two-stage analytical strategy was deliberate: ANOVA answers whether the group changed, while LGCM answers how individuals changed and what drove those differences. Institutional review board approval was obtained prior to data collection, and all participants provided informed written consent.

3.2. Participants

Participants were 55 first-years undergraduate EFL students enrolled in writing courses at Saudi universities, selected through purposive sampling. This type of sampling method was used to ensure that the participants were currently enrolled in an AI-integrated writing instruction. They were able to effectively interact with the survey tool throughout all three measurement waves. They were representative of the first-year undergraduate student population within the higher education sector in Saudi Arabia. Random sampling would not have guaranteed theoretical relevance to the research questions (Nyimbili & Nyimbili, 2024). Only students who completed all three measurement waves were retained in the final analysis. Attrition management was treated as a design priority rather than an afterthought, because incomplete longitudinal datasets introduce systematic bias — students who drop out mid-study are rarely a random subset, and their exclusion can distort both group means and growth estimates. Retaining only complete-case participants preserved the integrity of the developmental trajectories modelled in the LGCM. Table 1 presents the full demographic breakdown of the final sample.

Table 1: Demographic Characteristics of Participants (N = 55).

Characteristics	Categories	Frequency	Percent
Gender	Male	15	27.0%
	Female	40	73.0%

Restricted - مقيد

Age	22–30	7	12.73%
	30–35	31	56.36%
	35–40	17	30.91%
English Proficiency	Beginner	20	36.36%
	Intermediate	25	45.45%
	Advanced	5	9.09%
	Proficient	5	9.09%
Technology Proficiency	Not comfortable	12	21.82%
	Somewhat comfortable	28	50.91%
	Comfortable	10	18.18%
	Very comfortable	5	9.09%
Prior AI Tool Use	Yes	34	61.82%
	No	21	38.18%

3.3. Instrumentation

Data were collected using the AI Writing Engagement Scale (AI-WES), developed for this study by adapting the Academic Engagement Scale (Reeve & Tseng, 2011; Reeve, 2013). The original scale was designed for general academic contexts and required substantive modification to reflect the specific demands of AI-mediated writing tasks. This distinction matters: a student can be highly engaged in a writing class while remaining disengaged from the AI tools embedded within it. The adapted scale captures engagement with the AI-mediated writing task specifically, which is what the research questions require.

This final scale consists of 20 items, equally divided among four subscales, each of which is measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Representative items included: "I think deeply about how I can improve my writing with the help of an AI tool" (cognitive); "I actively participate in writing activities with the help of an AI tool" (behavioral); "I feel more motivated to write with the help of an AI tool" (emotional); and "I ask for more guidance if I do not know how to use an AI tool" (agentic).

Prior to data collection, all items were reviewed by a panel of university professors and experienced EFL instructors for clarity, cultural appropriateness, and theoretical alignment; several items were revised based on their feedback. Cronbach's alpha coefficients ranged from 0.810 to 0.855 across subscales and time points, confirming not only strong internal consistency but longitudinal measurement stability — a prerequisite for meaningful comparison of scores across waves.

Two open-ended reflective questions were administered alongside the survey at each time point to capture students perceived engagement changes and comparisons with prior writing experiences.

3.4. Intervention Procedure

The six-week intervention comprised three stages. At baseline (Week 1, Time 1), participants completed the engagement survey and reflective questions before structured AI tool use began, establishing a genuine pre-intervention benchmark. At the midpoint (Week 3, Time 2), the survey and questions were readministered following three weeks of integrated AI writing assistance. At post-intervention (Week 6, Time 3), participants completed the final survey and reflections upon course completion. AI tools were integrated into regular course writing tasks throughout the six weeks with instructor guidance.

The following section describes the procedures used to analyse the qualitative responses collected during the study.

3.5. Qualitative Analysis

Open-ended responses were analysed using thematic analysis following the six-phase approach described by Braun and Clarke (2006). In the first phase, all responses across all three time points were read repeatedly to achieve broad familiarity with the data before any coding began.

In the second phase, initial codes were generated inductively — short labels applied to meaningful units of text that captured how students described their engagement experiences, their perceived changes over time, and their comparisons with previous writing methods. These codes were developed without imposing the four-dimensional engagement framework in advance; the intention was to let the data surface its own patterns before testing whether they corresponded to the theoretical constructs.

In the third phase, codes were sorted into candidate themes grouping related ideas across participants and time points. In the fourth phase, themes were reviewed for internal coherence and distinctiveness, with several initial groupings merged or refined. The final theme set reflected how each engagement dimension — cognitive, behavioral, emotional, and agentic — evolved from

Restricted - مقيد

Time 1 through Time 3, as perceived and described by students themselves. These themes were used to contextualise and illustrate the quantitative trajectory data, providing a human narrative that enriches the statistical slopes and intercepts, rather than constituting an independent qualitative strand with its own claims.

3.6 Quantitative Analysis

Quantitative data were analysed using SPSS (v29) and AMOS 31, with the two packages serving complementary purposes. SPSS was used for descriptive statistics, normality testing, reliability analysis, and repeated-measures ANOVA; AMOS was used for the structural equation modelling framework within which LGCM was estimated. This dual-software approach is standard practice in longitudinal engagement research and allowed each analytical stage to draw on the most appropriate toolset (Grimm & Ram, 2018).

Before applying the repeated measures ANOVA, Mauchly's test indicated a significant violation of the sphericity assumption ($W = 0.048$, $p < .001$). Therefore, Greenhouse–Geisser corrections ($\varepsilon = 0.512$) were applied to adjust the degrees of freedom and produce conservative F-statistics.

Multiple normality tests were conducted and revealed significant non-normality across time points for all engagement dimension measures. Table 5 below illustrates the results from the normality tests conducted. The Shapiro-Wilk values obtained ranged from 0.770 to 0.830 and are below the 0.90 threshold commonly used to determine normality.

However, since the sample includes more than 50 participants, parametric tests such as ANOVA remain robust even when this assumption is violated because the effect of the violation is reduced by the central limit theorem. This violation is reported transparently rather than assumed, consistent with current reporting standards in applied linguistics and educational psychology.

The Latent Growth Curve Model was then applied and estimated with the help of the AMOS software. The intercept was set to have equal loadings across all time points, while the slope was set to have loadings of 0, 1, and 2 to represent Time 1, Time 2, and Time 3 respectively. The technological proficiency was entered as a time-invariant covariate and was set to have paths to the intercept and slope factors for each engagement dimension. The paths tested whether digital readiness predicted the level from which and how fast the student grew. The following section reports findings from each analytical stage in sequence, beginning with descriptive trends and progressing to the LGCM moderation results.

4. RESULTS

4.1. Descriptive Statistics

Table 2 presents descriptive statistics for all four engagement dimensions across the three measurement points. All dimensions increased substantially and consistently from Time 1 to Time 3, with emotional engagement reaching the highest end-point mean ($M = 4.000$) and also showing the greatest reduction in variability — suggesting that motivational benefits of AI tools became broadly shared across the group by post-intervention.

Table 2: Descriptive Statistics for Engagement Dimensions Across Three Time Points.

Dim.	N	T1 M	T1 SD	T2 M	T2 SD	T3 M	T3 SD	Range
CE	55	1.603	0.541	2.825	0.527	3.883	0.553	1–5
BE	55	1.625	0.643	2.495	0.752	3.905	0.697	1–5
EE	55	1.530	0.708	2.960	0.470	4.000	0.455	1–5
AE	55	1.538	0.779	2.654	0.690	3.665	0.582	1–5

Note: CE = Cognitive Engagement; BE = Behavioral Engagement; EE = Emotional Engagement; AE = Agentic Engagement. T1 = Time 1 (Baseline); T2 = Time 2 (Midpoint); T3 = Time 3 (Post-Intervention).

4.2. Normality And Reliability

Table 3 reports Shapiro-Wilk normality statistics and Cronbach's alpha values across all three time points. Normality was violated for all dimensions (all $p < .05$), with Shapiro-Wilk values ranging from 0.770 to 0.830. Despite this, parametric analyses were retained given the sample size above 50 and consistent with established guidance for moderate violations (Rovine & McDermott, 2018). Reliability was strong and stable, with alpha values ranging from 0.810 to 0.855, confirming measurement consistency across the longitudinal study.

Table 3: Shapiro-Wilk Normality Test and Cronbach's Alpha Reliability by Engagement Dimension.

Dim.	W	Skewness	Kurtosis	df	p	Cronbach's α (T1, T2, T3)
CE	0.780	-2.050	4.550	55	.000	0.850 / 0.848 / 0.855
BE	0.830	-1.730	3.350	55	.000	0.840 / 0.837 / 0.845
EE	0.770	-1.670	2.100	55	.000	0.800 / 0.810 / 0.815

Restricted - مقيد

AE	0.820	-1.910	5.240	55	.000	0.810 / 0.815 / 0.825

Note: W = Shapiro-Wilk statistic. All p-values < .05 indicate significant deviation from normality. Cronbach's α reported for Time 1, Time 2, and Time 3 respectively.

Table 4 confirms the sphericity violation that necessitated Greenhouse-Geisser corrections throughout the repeated-measures analyses.

Table 4: Mauchly's Test of Sphericity and Correction Epsilons.

Measure	Mauchly's W	Approx. χ^2	df	Sig.	G-G ε	H-F ε
CE, BE, EE, AE	0.048	160.870	2	< .001	0.512	0.513

Note: G-G = Greenhouse-Geisser; H-F = Huynh-Feldt. Greenhouse-Geisser correction (ε = 0.512) applied to all subsequent F-tests.

Table 5 presents results from three independent normality tests, all confirming significant non-normality across measurement points.

Table 5: Kolmogorov-Smirnov, Shapiro-Wilk, And Anderson-Darling Normality Tests Across Time Points.

Test	T1 Stat	T1 p	T2 Stat	T2 p	T3 Stat	T3 p
Kolmogorov-Smirnov (D)	0.161	.001	0.150	.003	0.144	.006
Shapiro-Wilk (W)	0.780	< .001	0.795	< .001	0.820	< .001
Anderson-Darling (A)	0.210	< .001	0.190	< .001	0.205	< .001

Note: All tests are reported for combined engagement dimensions. Significant p-values (< .05) across all three tests and all time points confirm non-normality.

4.3. Repeated-Measures Anova

Table 6 presents the repeated-measures ANOVA results. Time exerted a statistically significant effect on all four dimensions (all $p < .001$). F-values ranged from 296.67 (CE) to 548.42 (BE), and partial eta squared values from 0.918 to 0.954, indicating large and consistent effects across the study period.

Table 6: Repeated-Measures ANOVA Results for Engagement Dimensions Over Time.

Dimension	Value	F	Hyp. df	Err. df	Sig.	Partial η^2
Cognitive (CE)	0.918	296.672	2.000	53.000	< .001	0.918
Behavioral (BE)	0.954	548.417	2.000	53.000	< .001	0.954
Emotional (EE)	0.948	482.300	2.000	53.000	< .001	0.948
Agentic (AE)	0.933	370.901	2.000	53.000	< .001	0.933

Note: All F-tests use Greenhouse-Geisser correction. Partial η^2 values above 0.14 indicate large effects (Cohen, 1988).

Post-hoc pairwise comparisons (Table 7) confirmed that differences between all-time point pairs were significant for CE, BE, and EE. For AE, the T2 to T3 difference was non-significant ($p = .965$), suggesting that agentic growth was concentrated in the first half of the intervention.

Table 7: Post-Hoc Pairwise Comparisons for Engagement Dimensions (Bonferroni-Adjusted).

Dimension	(I)	(J)	Mean Diff (I−J)	SE	Sig.	95% CI
Cognitive (CE)	1	2	−2.222*	0.090	< .001	[−2.445, −1.998]
	1	3	−2.280*	0.094	< .001	[−2.512, −2.048]
	2	3	−0.058*	0.012	< .001	[−0.089, −0.028]
Behavioral (BE)	1	2	−1.869*	0.075	< .001	[−2.055, −1.683]
	1	3	−2.280*	0.069	< .001	[−2.450, −2.110]
	2	3	−0.411*	0.044	< .001	[−0.519, −0.303]
Emotional (EE)	1	2	−2.429*	0.082	< .001	[−2.632, −2.226]
	1	3	−2.469*	0.080	< .001	[−2.667, −2.272]
	2	3	−0.040*	0.014	.019	[−0.075, −0.005]
Agentic (AE)	1	2	−2.116*	0.078	< .001	[−2.309, −1.924]
	1	3	−2.127*	0.077	< .001	[−2.318, −1.936]
	2	3	−0.011	0.011	.965	[−0.038, 0.016]

Note: * p < .05. Mean differences are I minus J. All comparisons use Bonferroni adjustment for multiple testing.

4.4. Baseline Correlations and Moderation Effects

Table 8 presents zero-order correlations among engagement dimensions at baseline (Time 1). Behavioral and agentic engagement shared the strongest correlation ($r = .877$, $p < .01$), followed closely by cognitive-agentic ($r = .843$, $p < .01$) and cognitive-behavioral ($r = .805$, $p < .01$) links. Emotional engagement showed a notably weaker association with cognitive engagement ($r = .236$, ns), suggesting that initial emotional investment in AI tools did not automatically translate into deep cognitive processing at baseline — an important distinction for understanding how the four dimensions develop differentially over time.

Restricted - مقيد

Table 8: Zero-Order Correlations Among Engagement Dimensions at Baseline (Time 1).

	CE	BE	EE
Behavioral Engagement (BE)	0.805**	–	
Emotional Engagement (EE)	0.236	0.843**	–
Agentic Engagement (AE)	0.843**	0.877**	0.311*

Note: ** p < .01; * p < .05 (two-tailed). CE = Cognitive; BE = Behavioral; EE = Emotional; AE = Agentic.

Table 9 presents moderation regression results for the influence of technological proficiency. Technological proficiency was a significant direct predictor of cognitive (β = 0.35, $p < .001$) and behavioral engagement (β = 0.30, $p < .001$). Significant interaction effects further confirmed that proficiency moderated the relationship between engagement and growth over time (TP × CE: β = 0.12, $p = .017$; TP × BE: β = 0.10, $p = .014$; TP × Time: β = 0.08, $p = .045$).

Table 9: Moderation Regression: Influence of Technological Proficiency on Engagement Over Time.

Predictor	β	SE	t	p
Technological Proficiency (TP)	0.25	0.08	3.13	.002
Cognitive Engagement (CE)	0.35	0.07	4.99	< .001
Behavioral Engagement (BE)	0.30	0.06	5.00	< .001
TP × CE	0.12	0.05	2.40	.017
TP × BE	0.10	0.04	2.50	.014
TP × Time	0.08	0.04	2.00	.045

Note: TP = Technological Proficiency; CE = Cognitive Engagement; BE = Behavioral Engagement.

4.5. Latent Growth Curve Modelling

Table 10 presents LGCM results for all four engagement dimensions. All intercept and slope means were statistically significant, confirming genuine positive growth trajectories. Behavioral engagement showed the steepest slope (β = 0.45), followed by agentic (β = 0.41), cognitive (β = 0.38), and emotional engagement (β = 0.29). Significant variance around both intercepts and slopes confirmed meaningful individual differences — students varied not only in where they started but in how quickly and consistently, they grew across the six weeks.

Table 10: Latent Growth Curve Modelling Results for Engagement Dimensions.

Dim.	Intercept M	Slope M	Intercept σ^2	Slope σ^2	p (Intercept)	p (Slope)
CE	37.59	0.38	27.13	10.67	< .001	< .01
BE	34.53	0.45	25.83	8.35	< .001	< .01
EE	30.92	0.29	15.87	4.25	< .001	< .06
AE	33.18	0.41	23.46	9.58	< .001	< .01

Note: M = mean; σ^2 = variance. CE slope p < .01; BE slope p < .01; EE slope p < .06; AE slope p < .01.

4.6. Qualitative Evidence: Thematic Tracking Across Time Points

Table 11 summarises thematic shifts in students' open-ended reflections across the three measurement points, aligned with the corresponding quantitative trends. The qualitative patterns closely mirrored the statistical trajectories, providing a developmental narrative that enriches the LGCM output.

Table 11: Thematic Analysis of Open-Ended Reflections Across Three Measurement Points.

Dimension	Time 1 (Baseline)	Time 2 (Midpoint)	Time 3 (post-intervention)	Quantitative Alignment
Cognitive	Surface-level focus: "I mostly try to finish the task, but I'm not sure how to improve my ideas."	Strategy awareness: "AI helps me see where my ideas are weak, so I try to reorganize my writing."	Self-regulated thinking: "Now I plan my ideas, structure my paragraphs, and choose better words before I write."	Significant increase in CE means (T1=1.60 → T3=3.88)
Behavioral	Task compliance: "I only write what is required."	Increased effort: "I revise my writing more times when using AI tools."	Sustained practice: "I practice writing even outside of class using AI."	Strong repeated-measures effect (η^2=0.954)
Emotional	Anxiety: "Writing in English makes me nervous."	Reduced anxiety: "AI feedback makes me less afraid of mistakes."	Enjoyment: "I enjoy writing now and feel more confident."	EE increased steadily (T1=1.53 → T3=4.00)
Agentic	Passive: "I usually wait for the teacher's instructions."	Emerging initiative: "I ask questions about AI feedback."	Autonomy: "I decide when and how to use AI for my writing."	Strongest relative growth (slope=0.41)

Note: Representative student responses are paraphrased. Themes derived through inductive thematic analysis of responses at each time

Restricted - مقيد

point.

5. DISCUSSION

Taken together, the quantitative trajectories and qualitative reflections point to a coherent developmental pattern, the implications of which are examined below. The findings offer longitudinal evidence that AI writing tools, when integrated systematically over six weeks, are associated with sustained and progressive engagement growth across all four dimensions. That growth is not uniform — understanding who benefits most, and why, matters as much as acknowledging the overall upward trend.

5.1. Sustained Growth Across All Engagement Dimensions

RQ1 is answered unambiguously: all four dimensions improved significantly across the three measurement points, with large effect sizes (η^2 = 0.918–0.954). The LGCM results for RQ2 confirm that growth was linear and progressive — not concentrated in a single interval. The continuous rise from Time 2 to Time 3 (most notably in CE and BE, refer to Tables 6 and 7) contradicts the basic novelty effect explanation. If students were simply responding to the novelty of a new tool, engagement would have plateaued at the midpoint. Instead, it continued to accelerate, suggesting students were internalising the affordances of AI tools and developing more sophisticated engagement patterns as the course progressed (Luo & Yusuf, 2025; Shen et al., 2023).

The reduction in writing anxiety apparent in student reflections at Time 2 — shifting from "Writing in English makes me nervous" to "AI feedback makes me less afraid of mistakes" — provides a plausible affective mechanism for this trajectory. Lower anxiety creates space for more ambitious cognitive and strategic engagement. This is in line with dynamic systems theory's predictions that positive early conditions lead to self-reinforcing developmental cycles (López-Pernas & Saqr, 2023). This is further supported by the emotional engagement data showing the greatest absolute increase and the narrowest spread by Time 3 (M = 4.00, SD = 0.455), suggesting that affective benefits were democratically distributed — even students who entered the course anxious and uncertain had largely overcome that barrier by post-intervention.

5.2. Cognitive And Agentic Engagement as the Core Developmental Cluster

Among the four dimensions, cognitive and agentic engagement exhibited the strongest baseline correlation (r = .843, Table 8) and showed substantial slope growth (Table 10). This co-development reflects what Reeve and Tseng (2011) identify as the highest forms of engagement — not surface participation but deep thinking and active self-direction. Students at Time 3 who described deliberately deciding which AI suggestions to accept, self-generating revision criteria, and practising independently were demonstrating exactly this agentic orientation.

What is particularly significant is that AI tools functioned as cognitive scaffolds rather than cognitive replacements. Students described using AI to test their own ideas, compare alternative structures, and evaluate feedback — all of which are higher-order activities (Shen et al., 2023; Yan & Zhang, 2024). This corresponds with sociocultural theory's characterization of technology as a mediating instrument that enhances rather than replaces learner cognition (Lantolf & Thorne, 2006). The agentic growth slope (β = 0.41) and its continued rise through Time 3 indicates that the transition from dependent to strategic AI use is not immediate; rather, it requires time and repeated engagement, demonstrating the importance of continuous longitudinal instruction over temporary interventions.

5.3. Technological Proficiency as an Amplifier of Growth

The question of who grew most — and why — brings us to the study's third research question. Technological proficiency predicted both intercepts and slopes across dimensions (Table 9). Students with higher digital readiness began the course at higher engagement levels and grew faster — an amplification effect that widened the gap between high- and low-proficiency learners over the six weeks. The finding has a significant practical impact: without intentional intervention, AI-assisted education may accidentally worsen existing inequalities. Students with lower proficiency could face a different type of experience that features passive reliance on automated corrections instead of strategic revision, aligning with the dependency pattern that Xie et al. (2023) advise against.

Notably, emotional engagement showed weaker moderation by technological proficiency, indicating that the anxiety-reducing benefits of AI feedback were more broadly accessible regardless of digital skill level. By Time 3, even students who reported low technological comfort described feeling less

Restricted - مقيد

anxious and more confident. This is an encouraging finding: the affective floor-raising effect of AI tools appears to be inclusive even where cognitive and agentic benefits are not yet equally distributed. The overall amplification pattern means digital literacy preparation must accompany AI deployment — not as an afterthought, but as a prerequisite for equitable participation in AI-supported writing curricula (Zhou & Wei, 2018).

5.4. Implications For AI-Supported Writing Pedagogy

The findings carry direct and practical implications for the integration of AI tools into EFL writing pedagogy. Engagement does not improve simply because AI tools are present in a classroom. It improves when those tools are embedded within a coherent pedagogical framework that gives students the guidance, structure, and time they need to move from passive interaction with automated feedback to active, strategic use of it. Three interrelated implications follow from the evidence.

First, AI tools should be seen as educational tools that need careful planning, not as self-explanatory resources that operate independently of teaching. The longitudinal growth observed across all four engagement dimensions in this study occurred within a structured six-week instructional context, where the use of AI tools was facilitated by instructors and integrated into regular writing tasks. This aligns with Luo and Yusuf's (2025) conclusion that AI-adaptive feedback most effectively translates writing engagement when accompanied by pedagogical scaffolding that aids students in interpreting and responding to the feedback. Without that support, students, particularly those with limited technological proficiency, risk engaging with AI output at a surface level, accepting suggestions without evaluating them and missing the deeper cognitive and metacognitive benefits the tools are capable of supporting. Curriculum designers and writing instructors should therefore plan explicitly for how AI feedback will be introduced, modelled, discussed, and progressively handed over to students across an instructional sequence, rather than assuming that access to the tool is sufficient on its own.

Second, digital literacy development must be treated as a prerequisite for AI-supported writing instruction rather than an optional supplement. The moderation analysis in this study showed clearly that technological proficiency amplified the benefits of AI integration — students who arrived with stronger digital skills grew faster and engaged more strategically. This widening gap is only a problem if institutions deploy AI tools without attending to digital readiness. For first-year undergraduates in Saudi universities, where technological comfort levels vary substantially as confirmed by the demographic data here, this means building foundational digital skills into early writing courses — not as a separate IT module, but as an integrated component of writing pedagogy. Students need to learn not just how to use an AI writing tool operationally, but how to read its suggestions critically, understand their limitations, and make informed decisions about what to accept, adapt, or reject. These are interpretive and evaluative capacities, and they are as much a part of academic writing development as grammar or organisation (Fredricks et al., 2004; Zhou & Wei, 2018).

Third, instructors have a specific and irreplaceable role to play in scaffolding strategic AI use over time. The qualitative data in this study showed that the shift from passive task-completion to agentic, self-directed writing did not happen immediately — it developed gradually across the six weeks, with students describing increasingly deliberate and independent choices about how to engage with AI feedback at Time 2 and Time 3. This developmental trajectory exemplifies the process-to-product transition articulated by Shen et al. (2023): as students receive iterative AI feedback and are encouraged to contemplate it, they progressively assimilate strategic revision practices that endure beyond specific writing assignments. Teachers can help make this change happen by showing students how to evaluate AI suggestions, making assignments that make students explain why they made changes, and including structured reflection activities, like the open-ended questions used in this study, in the teaching cycle. The goal is not for students to use AI tools more, but to use them better: with intention, critical judgment, and a clear sense of their own authorial goals. When instructors scaffold this kind of strategic engagement from the beginning of a course, AI tools become genuine developmental supports rather than convenient shortcuts.

5.5. Limitations

These pedagogical possibilities must be weighed against the study's methodological boundaries. The sample of 55, while producing stable LGCM estimates, falls below recommended thresholds for SEM-based analyses, and parameter estimates warrant cautious interpretation. The absence of a control group means engagement growth cannot be attributed exclusively to AI integration — instructor

Restricted - مقيد

effects, course progression, and repeated assessment remain plausible contributing factors. The six-week duration, while sufficient to observe linear growth, limits claims about long-term sustainability. Future research should replicate this design with larger samples, include comparison conditions, specify AI tool use precisely, and extend the observation period to capture potential plateaus or reversals. With those limitations in view, the following conclusion draws together the study's central contributions.

6. CONCLUSION

This study set out to examine whether AI writing tools produce lasting engagement growth or temporary enthusiasm among first-year EFL learners. Within the constraints of this six-week longitudinal design, the answer is clear: meaningful, progressive growth is possible across all four dimensions of engagement — and it is most pronounced in the dimensions that matter most educationally, cognitive depth and learner agency. The quantitative trajectories confirmed by LGCM and the developmental narratives captured in student responses (Table 11) together tell a coherent story: AI-mediated writing environments can transform first-year learners from anxious, task-compliant writers into reflective, self-directed ones — provided students develop the necessary digital skills to engage strategically rather than passively.

For Saudi higher education institutions integrating AI into writing curricula, the practical implication is direct: deploy AI tools alongside structured digital literacy preparation, and use longitudinal assessment frameworks that can capture developmental change rather than snapshot performance. AI tools have genuine potential to support equitable and transformative writing development — but realising that potential requires treating technological competence as a curriculum priority, not an assumption.

REFERENCES

Bergdahl, N., Nouri, J., & Fors, U. (2020). Disengagement, engagement and digital skills in technology-enhanced learning. Education and Information Technologies, 25(2), 957–983. https://doi.org/10.1007/S10639-019-09998-W

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109. https://doi.org/10.3102/00346543074001059

Grimm, K. J., & Ram, N. (2018). Latent growth and dynamic structural equation models. Annual Review of Clinical Psychology, 14(1), 55–89. https://doi.org/10.1146/ANNUREV-CLINPSY-050817-084840

Haji, N. (2022). Major writing difficulties faced by Kurdish learners of EFL in academic writing classes. Twejer. https://doi.org/10.31918/twejer.2251.27

Henrie, C. R., Bodily, R., Manwaring, K. C., & Graham, C. R. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. International Review of Research in Open and Distributed Learning, 16(3), 131–155. https://doi.org/10.19173/irrodl.v16i3.2015

Hubbard, P. (2013). Making a case for learner training in technology enhanced language learning environments. The CALICO Journal, 30(2), 163–178. https://doi.org/10.11139/CJ.30.2.163-178

Kim, M. K., Lee, I. H., & Kim, S. M. (2021). A longitudinal examination of temporal and iterative relationships among learner engagement dimensions during online discussion. Technology, Knowledge and Learning, 8(1), 63–86. https://doi.org/10.1007/S40692-020-00171-8

Lai, Z. C.-C. (2025). The impact of AI-assisted blended learning on writing efficacy and resilience. International Journal of Computer-Assisted Language Learning and Teaching, 15(1). https://doi.org/10.4018/IJCALLT.20250101

Lantolf, J. P., & Thorne, S. L. (2006). Sociocultural theory and the genesis of second language development. Oxford University Press.

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. Social Behavior and Personality, 46(3), 517–528. https://doi.org/10.2224/SBP.7054

Li, J., & Fan, Y. (2024). Does engagement with feedback matter? Unveiling the impact of learner engagement and grit on EFL learners' English writing achievements. Language Teaching Research. https://doi.org/10.1177/13621688241257865

Liu, S., Li, S., & Xing, J. (2025). Exploring Chinese secondary EFL students' self-regulated learning and task engagement in AI-assisted classrooms: A latent growth curve modelling study. European Journal of Education, 60(4). https://doi.org/10.1111/ejed.70241

López-Pernas, S., & Saqr, M. (2023). How the dynamics of engagement explain the momentum of achievement and the inertia of disengagement: A complex systems theory approach. Computers in Human Behavior,

Restricted - مقيد

153, 108126. https://doi.org/10.1016/j.chb.2023.108126

Luo, L., & Yusuf, A. (2025). Bridging AI and pedagogy: How AI-adaptive feedback shapes Chinese EFL students' writing engagement, metacognitive writing strategies, and writing performance. Assessment & Evaluation in Higher Education, 1–24. https://doi.org/10.1080/02602938.2025.2548919

Oktarina, I. B., Saputri, M. E. E., Magdalena, B., Hastomo, T., & Maximilian, A. (2024). Leveraging ChatGPT to enhance students' writing skills, engagement, and feedback literacy. Edelweiss Applied Science and Technology, 8(4), 2306–2319. https://doi.org/10.55214/25768484.v8i4.1600

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. Journal of Educational Psychology, 105(3), 579–595. https://doi.org/10.1037/a0032690

Reeve, J., & Tseng, C. (2011). Agency as a fourth aspect of students' engagement during learning activities. Contemporary Educational Psychology, 36, 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Reeve, J., Cheon, S. H., & Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. Contemporary Educational Psychology, 62, 101899. https://doi.org/10.1016/j.cedpsych.2020.101899

Rovine, M. J., & McDermott, P. A. (2018). Latent growth curve and repeated measures ANOVA contrasts: What the models are telling you. Multivariate Behavioral Research, 53(1), 90–101. https://doi.org/10.1080/00273171.2017.1387511

Shen, C., Shi, P., Guo, J., Xu, S., & Tian, J. (2023). From process to product: Writing engagement and performance of EFL learners under computer-generated feedback instruction. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2023.1258286

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wong, Z. Y., & Liem, G. A. D. (2021). Student engagement: Current state of the construct, conceptual refinement, and future research directions. Educational Psychology Review, 1–32. https://doi.org/10.1007/S10648-021-09628-3

Xie, Y., Jang, J. R., Kim, S., & Cho, Y. H. (2023). Interaction patterns between learners and AI tools for English writing. Proceedings of ICCE 2023. https://doi.org/10.58459/icce.2023.1500

Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. Humanities and Social Sciences Communications, 11, 1086. https://doi.org/10.1057/s41599-024-03543-y

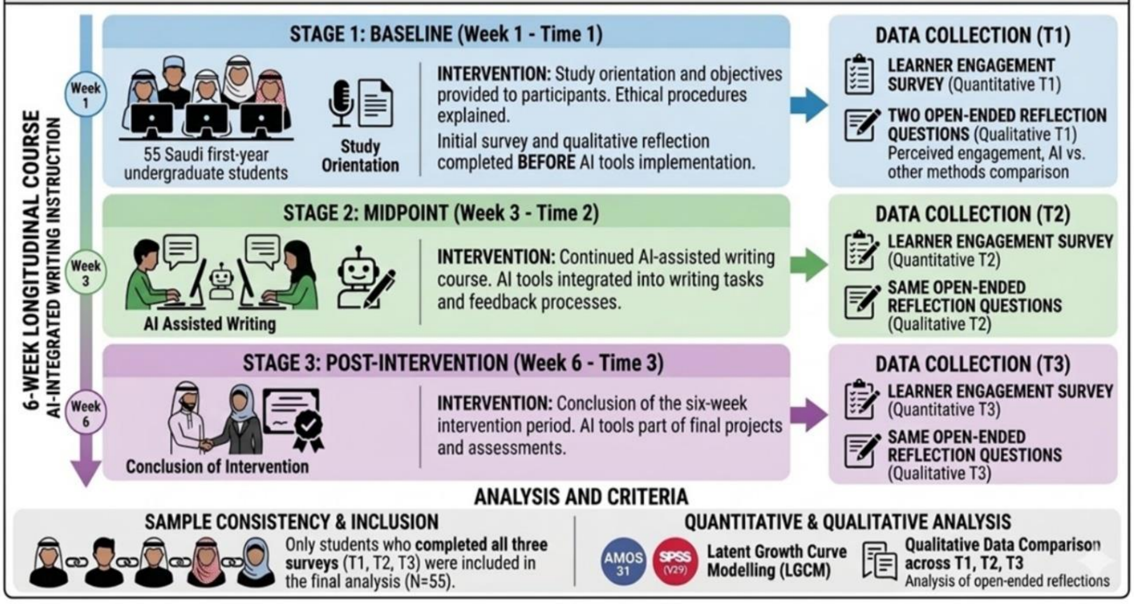
Zhang, F. (2022). Latent growth curve modeling for the investigation of emotional factors in L2 in longitudinal studies: A conceptual review. Frontiers in Psychology, 13, 1005223. https://doi.org/10.3389/fpsyg.2022.1005223

Zhou, S., Hiver, P., & Zheng, Y. (2023). Modeling intra- and inter-individual changes in L2 classroom engagement. Applied Linguistics, 44(6), 1047–1076. https://doi.org/10.1093/applin/amac065

Zhou, Y., & Wei, M. (2018). Strategies in technology-enhanced language learning. Studies in Second Language Learning and Teaching, 8(2), 471–495. https://doi.org/10.14746/SSLLT.2018.8.2.13

Restricted - مقيد

FIGURE X: STUDY INTERVENTION PLAN AND LONGITUDINAL DATA COLLECTION FLOW (AI-ASSISTED WRITING STUDY)



Restricted - مقيد