



## Research article

## Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial

Nabi Nazari<sup>a,\*</sup>, Muhammad Salman Shabbir<sup>b</sup>, Roy Setiawan<sup>c</sup><sup>a</sup> Department of Psychology, Faculty of Human Sciences, Lorestan University, Khorramabad, Iran<sup>b</sup> Department of Management, College of Commerce and Business Administration, Dhofar University, 211 Salalah, Oman<sup>c</sup> Universitas Kristen Petra, Indonesia

## ARTICLE INFO

## Keywords:

Artificial Intelligence  
Automated writing evaluation  
L2 writing  
Feedback  
Pedagogy  
ESL  
Formative assessment

## ABSTRACT

A major challenge in educational technology integration is to engage students with different affective characteristics. Also, how technology shapes attitude and learning behavior is still lacking. Findings from educational psychology and learning sciences have gained less traction in research. The present study was conducted to examine the efficacy of a group format of an Artificial Intelligence (AI) powered writing tool for English second postgraduate students in the English academic writing context. In the present study, (N = 120) students were randomly allocated to either the equipped AI (n = 60) or non-equipped AI (NEAI). The results of the parametric test of analyzing of covariance revealed that at post-intervention, students who participated in the AI intervention group demonstrated statistically significant improvement in the scores, of the behavioral engagement (Cohen's  $d = .75$ , 95% CI [0.38, 1.12]), of the emotional engagement Cohen's  $d = .82$ , 95% CI [0.45, 1.25], of the cognitive engagement, Cohen's  $d = .39$ , 95% CI [0.04, .76], of the self-efficacy for writing, Cohen's  $d = .54$ , 95% CI [0.18, 0.91], of the positive emotions Cohen's  $d = .44$ , 95% CI [0.08, 0.80], and of the negative emotions, Cohen's  $d = -.98$ , 95% CI [-1.36, -0.60], compared with NEAI. The results suggest that AI-powered writing tools could be an efficient tool to promote learning behavior and attitudinal technology acceptance through formative feedback and assessment for non-native postgraduate students in English academic writing.

## 1. Introduction

Academic writing can be a problematic, affective, and complex process (Rahimi and Zhang, 2018) that has a prominent scientific career. Academic writing in the English language is a sophisticated, essential, and integrative task that is difficult for native and international students (Campbell, 2019). Additionally, for English second learners (ESL), the difficulty is combined with linguistic and educational barriers (Hanauer et al., 2019). Unfortunately, very little is being done in higher education to equip graduate students. Monitoring the writing process and offering valuable productive feedback to students is too time-consuming, labor-intensive, and subjective (Lim and Phua, 2019). The computer-based applications are increasingly becoming alternatives to facilitate writing using automated writing evaluation (AWE), automated essay scoring (AES), and automated written corrective feedback (AWCF). New writing tools, powered by Artificial Intelligence (AI) and available in mobile devices, are promising tools to assist students in learning and develop writing skills that are hard to learn from traditional training.

AI is an umbrella term to describe an automated device that can behave of human intelligence processes such as learning, reasoning, and self-correction (Popenici and Kerr, 2017). One of the most important goals for AI is to design automated devices that can analyze the environment and do a task as humans do. New writing applications potentially offer flexible and time-saving additions to the writing curriculum are integrated to provide the AWE, AES, and AWCF features in one integrated application (Koltovskaia, 2020). With advances in technology, AI new teaching and learning experiences in the assessment, tutoring, content generation, and feedback for teachers and students. Perhaps the most contributions of digital writing tools are defined throughout the formative feedback and assessment. Moreover, the new AI Applications provide a comprehensive instructional practice and plagiarism detection component that may assist ESL in research writing progress (Zawacki-Richter et al., 2019). Furthermore, the idea of AI compounded with Mobile teaching and learning (m-learning) is emerging in higher education (Pedro et al., 2018), which can afford new opportunities to

\* Corresponding author.

E-mail address: [Nazariirani@gmail.com](mailto:Nazariirani@gmail.com) (N. Nazari).

enhance pedagogical flexibility, learning process or outcome, and feedback immediacy (Cheung, 2015).

### 1.1. Formative feedback

Students in traditional educational settings rarely receive the preferred or required formative feedback (Molin et al., 2020). Instructors no longer have sufficient time to provide formative feedback comments due to overloaded tasks or overcrowded classrooms. Also, without scoring, teachers may lack the ability to assess students' progress. Some researchers have even ex-teachers' writing proficiency has been questioned. L2 writing teachers have few professional development opportunities due to heavy workloads and limited resources (Yu et al., 2019; Hegelheimer and Lee, 2013). For research writing, learners must be self-regulated, responsible, intrinsically motivated, and engaged actively to deep and meaningful learning. From a Constructivist viewpoint, deep and meaningful learning can be defined as "an active, constructive, self-regulated, cumulative, goal-oriented, collaborative and individual process of knowledge-building and construction of meaning based on prior knowledge and placed in a specific context" (Jonassen et al., 2003). In this sense, one of the most important contributions of AI in education and meaningful learning is giving immediate feedback to the students about the learning progress. To some extent, feedback delayed is feedback denied (Wible et al., 2001). To foster student engagement, improve achievement, motivation, and self-regulation (Zimmerman and Labuhn, 2012), formative feedback needs to be immediate; online learning environments support and facilitate this immediacy. The feedback allowed students to become more engaged, knowledge builders, active and autonomous. The real-time formative feedback, followed by practical and instructional examples, provides new possibilities for learners' more personalized experiences. There is evidence that instructional feedbacks may enhance writing problem-solving (Li et al., 2014) and self-regulatory strategies (Wang et al., 2013). Regarding the high workload of thesis advisors, AI feedbacks can play an increasingly valuable role in L2 writing (Zhang, 2020). New AI-powered writing tools provide the reliable and accurate formative and summative assessment to provide information about learning. As an authentic pedagogical practice students' self-evaluation alludes to the tendency and capability to precisely assess and monitor student's knowledge while learning (Fook, 2010). Authentic pedagogy is associated with "Constructivism," is leading to self-development and focusing on learning rather than grading or memorization. Through this self-evaluation process, the writer improves the final written drafts and revising skills (Cavaleri and Dianati, 2016; Fadhilah et al., 2019; Parra and Calero, 2019).

### 1.2. Current issues

The use of automated systems to assess student progress and provide feedback is increasing for writing. With advances in technology, AI provides new teaching and learning experiences in the assessment, tutoring, content generation, and feedback for teachers and students. While AI is a growing field in education, AI application in L2 writing brings about several issues. Studies on digital writing applications in L2 writing contexts remain few. Relatively, findings from educational psychology and other learning sciences have gained less traction in research. However, studies have recognized the AI effect on writing outcomes, how AIs shape learning behavior is still lacking. Preliminary research supports the effectiveness of the several technologies to improve student engagement; few studies used randomized controlled experimental designs; therefore, very little evidence exists to substantiate a cause and effect relationship between technologies and student engagement indicators (Schindler et al., 2017). The impact of AI on writing is understudied, particularly for grammatical feedback. L2 writers concentrate overwhelmingly on grammatical errors, which may overlook the other aspects of research writing. However, technological advancements in AI writing are extended. Affective provisions, motivational features,

in-depth analytic learning, plagiarism detector, and social interaction platforms may promote writing skills, engagement, and non-cognitive traits. However, how best to utilize the portable tools to foster student development in academic skills and well-being is a little-studied field (Voogt et al., 2018). A major challenge in Educational Technology Integration is to engage students with different affective characteristics, motivation, beliefs, self-concepts, and emotions. To find answers for how technology shapes attitude and learning behavior through technology formative feedback, the current study focused on the engagement, self-efficacy, and academic emotions, which are representative constructs in educational psychology to address learning behavior and attitude.

### 1.3. The experimental framework

*Engagement:* Students' engagement is a critical factor in addressing low achievement problems, emotional state, course retention, or dropout rates (Fredricks and McColskey, 2012). According to Fredricks et al. (2004), student engagement can be conceptualized in three dynamically interconnected dimensions. Behavioral engagement encompasses students' participation ideas in academic or extracurricular activities and to make an effort to learn tasks. Emotional engagement refers to positive (e.g., enjoyment) or negative (e.g., anxiety) affective reaction to the learning environment. Emotions play a central role in scientific models of human development, interpersonal processes, learning, and decision-making (Lobue et al., 2019; Nazari and Griffiths, 2020). In this sense, emotional engagement is recognized as the central factor of engagement. Cognitive engagement refers to time investment or intrinsically motivated efforts and set the learning strategy to mastery. While the growing technology offers numerous learning opportunities, Regardless of the tech-powered innovation level, meaningful learning depends on how students are actively engaged in the specific task. Preliminary support for purposefully implemented technology may positively influence student engagement (Wang et al., 2019; Schindler et al., 2017). However, in web-based learning systems, students' degree of engagement in educational learning is lower than in face-to-face education systems (Skinner et al., 2014). Students are self-regulated to the degree that they are met cognitively, motivationally, and behaviorally active participants in their learning process (e.g., Pintrich, 2000). The integration of behavioral, emotional, and cognitive dimensions roughly covers some of the most-researched areas in language learning research, such as motivation, affective orientations, cognitive traits, and learning strategies.

A major challenge in Educational Technology Integration is to engage students with different affective characteristics, motivation, beliefs, self-concepts, and emotions. Specifically, Attitudinal Acceptance is conceptualized to address challenging individual differences related to technology effectiveness and integration (Kopp et al., 2008). Attitudinal acceptance refers to affective and cognitive experiences with the technology environment during learning (Mandl and Bürg, 2008). While cognitive acceptance relates to reason based component comparison of the technology offering efficacy and cost-effectiveness, Affective Acceptance considers as feeling-based component (e.g., enjoyment, anxiety). Self-efficacy (subjective and judgmental beliefs) (Brinkerhoff, 2006) and academic emotions during learning are considered representative constructs to address affective acceptance. Also, higher self-efficacy levels are associated with more heightened experienced positive emotions.

*Self-efficacy:* The first factor in determining behavior is how an individual has a positive self-efficacy for acquiring a behavior. Self-efficacy is a well-documented predictor of behavior (Bandura, 1977). Self-efficacy is defined as the expression of personal beliefs related to their capability to succeed in a specific behavior or to learn or perform a particular task (Zimmerman and Kitsantas, 2007) effectively. Efficacy beliefs are directly associated with motivation (Bandura, 1977, 2015), intrinsic motivation (Ryan and Deci, 2018),

and professional commitment. Interestingly, Self-efficacy can affect performance via goal orientation, and performance impacts self-efficacy again (Nazari, 2020; Lishinski et al., 2016). Writing self-efficacy refers to a students' feelings of competence as a writer (Bruning et al., 2013) or individual confidence in writing (Christensen and Knezek, 2017). Self-efficacy beliefs predict EFL learners' English proficiency (e.g., Wang and Bai, 2017). In the technology acceptance model (Dumbauld et al., 2014), self-efficacy is one of the main factors associated with higher acceptance. *Self-efficacy* is an affective construct that is associated with motivation and learning effectiveness. Broadly, research has evaluated writing-related skills so that research about writers' self-efficacy has received less attention (Bruning et al., 2013). These relationships are consistent with Social Cognitive Theory, which supports the idea that confident individuals believe their actions will produce successful results (Pajares and Schunk, 2002). The self-efficacy for writing model, proposed by (SEWS, Bruning et al., 2013), operationalizes Self-efficacy for writing in three domains. Self-efficacy (SE) for ideation (i.e., self-beliefs about the ability to generate ideas), *SE for conventions* (i.e., self-beliefs about adhering to language rules), and *SE for regulation* (i.e., self-beliefs about regulating writing behavior). For writing, Self-regulatory skills are critical to product creation ideas, expressing ideas, writing strategies, and writing willingness (Wei and Chou, 2020).

*Emotions:* Writing is an affective experience. Research has well-documented a link between learning affective dimensions and achievement in traditional education. University students experience a range of positive emotions (e.g., enjoyment, pride) and negative emotions (e.g., anxiety, boredom) in academic settings. Academic

emotions affect goal-orientation (Pekrun et al., 2009), self-regulated strategy, behavioral control, and academic success (Nazari and Far, 2019). Foreign language anxiety (FL anxiety) has long been recognized as a second language barrier (Sadiq, 2017). With the focus on the learner, rather than the learning outcomes, the knowledge about technology's impact on the learners' overall well-being (psychological and emotional) is less. Some studies argue that an online education environment may increase negative emotions. Lack of social interaction and lack of digital literacy may generate anxiety and boredom (Wosnitza and Volet, 2005). Collectively, these negative emotions appear to be particularly strong for L2 students engaged with technology. Importantly, concerning academic emotion transitions, there is evidence that students often remain in the same emotional condition under no intervention condition (Graesser, D'Mello and Strain, 2015).

1.4. The present study

A randomized controlled trial was conducted to examine the efficacy of a group format of an AI-powered writing tool for English second postgraduate students in the English academic writing context. We hypothesized that students who participate in the AI intervention group demonstrate significant improvement in engagement (behavioral, emotional, and cognitive) compared with the non-equipped AI (NEAI) group at post-intervention. We hypothesized that students who participate in the AI intervention group demonstrate significant improvement in individual acceptance factors (self-efficacy scores and academic emotions scores) compared with the non-equipped AI (NEAI) group post-intervention.

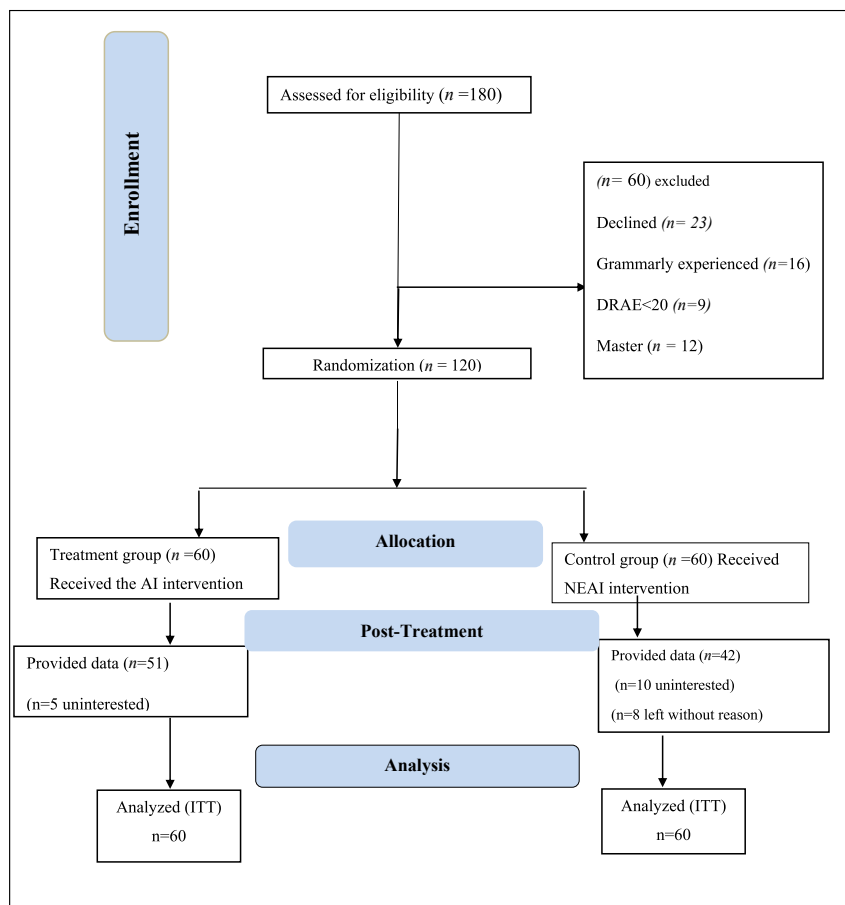


Figure 1. The participants flow chart diagram.

## 2. Method

### 2.1. Participants

The consort diagram is illustrated in Figure 1. Participant recruitment efforts included sending an e-mail to staff and advertisement in forums. A total of 180 students were assessed to participate in the trial. Of these, 120 (53 females, 67 males) met all eligibility criteria. Participants' ages ranged from 26 to 39 years. The mean age of the participants was 32.04 years ( $SD = 3.35$ ). The demographic characteristics of the sample are shown in Table 1. The inclusion criteria were: age over 18 years, being a postgraduate student of a national university (free, full-time), and acceptable digital literacy.

### 2.2. Instrument

**Grammarly:** Grammarly is an AI-powered product (Taguma et al., 2018) that offers an AWE, AES, and AWCF application in one digital writing tool with more than 20 million international users. Grammarly is also available for mobile devices (tablets, smartphones) with different operating systems. Also, Grammarly supports computer and mobile internet browsers, e-mail smartphone applications, and social media platforms, automatically. Grammarly is offered in two versions free version and the premium version. The premium version was used in the study. Although the free version provides feedback, the plagiarism detector and corrections are not covered in the free version.

**Digital Readiness (DRAE; Hong and Kim, 2018).** The DRAE, a 17-item, self-report measure, was used to assess the level of digital readiness and Students' perceived digital competencies for academic engagement five sub-scale. The students indicated the level of their agreement with five statements (1 = *strongly disagree*, 5 = *strongly agree*). The scores indicate a level of digital readiness and Students' perceived digital competencies for academic engagement. The scale shows strong reliability in this study ( $\alpha = .91$ ).

**Student Engagement scale. (Fredricks et al., 2005).** The student engagement scale is a 19-item self-report measure that determines to measure the students' engagement in three subscales. The five items related to behavioral (e.g., "I will conform to this online course's regulations") were used in our study. The six items were used to assess emotions (e.g., "I like taking this online course"), and eight-items were used to assess cognitive engagement (e.g., "I will find a way to comprehend the content of this online course when I cannot understand it"). The five-point Likert scale was employed (1 = *strongly disagree*; 5 = *strongly agree*).

Cronbach's value for behavioral, emotional, and cognitive engagement was .81, .77, and .79.

**Self-Efficacy for Writing Scale (SEWS; Bruning et al., 2013).** The SEWS, a 16-item self-report measure, was used to assess self-efficacy for writing. The SEWS consists of a 5-item for idea generation, a 5-item for the conventions, and a 6-item for self-regulation. The students indicated the level of their agreement with five statements (1 = *strongly disagree*, 5 = *strongly agree*). Internal consistency in the current study was acceptable ( $\alpha = 0.81$ ).

**Achievement emotion questionnaire (AEQ; Pekrun et al., 2011).** The AEQ was used to measure negative and positive emotions. The scale assesses emotion in three situations. In the study, the students only rated their experienced emotions during learning of English research in two subscales. For positive emotion, the 3-items related to enjoyment (e.g., "I enjoy acquiring new knowledge about writing research in English.") and for negative emotions, the 4-items related to anxiety (e.g., "I get tense and nervous while writing a research in English.") were employed to assess academic emotions on a 5-point Likert scale (1 = *completely disagree*, 5 = *completely agree*). Internal consistency in the current study was acceptable ( $\alpha = 0.83$ ).

### 2.3. Procedure

The study was a single-blind, parallel randomized controlled trial comparing the educational intervention group, based on the AI, with an NEAI control group. The study, including all assessments and interventions, was conducted via the internet. The IRB reviewed the research protocol to ensure ethical considerations, participant confidentiality, sampling, and obtaining informed consent.

**Sampling:** A multi-modal targeting strategy (McRobert et al., 2018) was used in three stages: (i) the most popular online communities for research were identified (ii) the forums had a user's public lists, and potential participants were invited via text message, and finally (iii), a generic online invitation letter was sent to all users. A public invitation letter was also posted on each community's public wall, seen by all users, periodically, on Monday, weekly. Over 40% of students were recruited in the first five weeks, 70% by Week 10, and the rest by Week 14.

First, interested postgraduate students were notified about the study's goals, session numbers, randomization, and group allocation chance through telephone, chats, and e-mail. Only those who provide a signed consent form to participate in the study; requested to complete the assessment protocol. To ensure inclusion criteria, the participants could send from verified academic email. At last, only consented subjects who

**Table 1.** Demographic characteristics and descriptive statistics of the sample (N = 120).

Item (N = 120)	Value	Test	P
<b>Gender, n (%)</b>			
Women	59 (48)	$\chi^2 = .04$	.85
Man	61 (52)		
<b>Field, n (%)</b>			
Humanity	29 (24.2)	$\chi^2 = 5.15$	.076
Technology	42 (35.0)		
Health	49 (40.8)		
<b>Continuous variables M (SD)</b>			
Age, years	32.04 (3.35)	$t(118) = .63$	.53
DRAE	32.53 (5.47)	$t(118) = -1.63$	.11
Self-efficacy	29.97 (4.71)	$t(118) = -1.95$	.054
Engagement	12.93 (6.53)	$t(118) = .12$	.91
Positive emotion	2.97 (.74)	$t(118) = -.39$	.69
Negative emotion	2.83 (.51)	$t(118) = .99$	.32

Note: n: frequency, M: Mean, SD: Standard Deviation.

t: independent t test between groups.

DRAE: Digital Readiness for Academic Engagement.



met inclusion criteria were selected for randomization. The outcomes were assessed at two time-points: Time 1: pre-intervention to pre-allocation includes baseline, Time 2: immediate after intervention: post-intervention assessment.

#### 2.4. Design

Student experiences with e-learning systems can affect academic engagement. The prior skills and background knowledge, English language ability, stereotype (gender; field) were controlled in three steps. *First:* the DRAE was used in enrollment to control digital experience. The participants must obtain a minimum score in DRAE >35. *Second:* to enroll inexperienced students in *Grammarly*, we did not address *Grammarly's* public plain announcement. The course was introduced as an online course for research writing. Also, a question with multiple choice was requested in the demographic characteristics of the software experience. The question options included the report or research writing software (e.g., word, Mendeley, excel, Grammarly). If Grammarly was checked, the participant was not included in the study. *Third:* we enrolled students in national universities. Students at these universities must obtain a minimum English score (equals 450 TOEFL Paper, 4.5 IELTS). Otherwise, they are not allowed to defend their dissertation. Moreover, the English language and research knowledge are the most determinant factors in the Ph.D. entrance exam. *Finally,* the allocation was carried out using a block size method, stratified by gender and field of the study.

##### 2.4.1. Sample size

The Sample size for analyzing covariance (ANCOVA) was conducted using G\*Power 3.1 analysis (Faul et al., 2009). A priori power analysis was conducted, using an alpha of .05, a power of 0.95, and a moderate to large effect size ( $f = .35$ ) to determine the sample size. According to G\*Power, the desired total sample size was 109. Therefore, 120 participants recruited, allowing for a 10% loss of data (dropping out prior intervention, post-intervention assessment).

##### 2.4.2. Randomization and blinding procedures

Randomization was performed using the block size stratified method. The block size was 6, stratified by gender and field. An independent statistician carried out the randomization and informed the participants and research team members about the allocation. The concealed was disclosed at the end of the study. To answer the question related to *Grammarly*, the first author was available via social media. The instructors (two associate professors) were not informed about the groups and aim of the study. To masking condition assisting, participants were instructed not to disclose any information about the intervention status. Evaluators, instructors, assessors, and statistic investigators were blinded to the intervention, participants' group.

##### 2.4.3. Experimental manipulation

**2.4.3.1. AI intervention.** The AI course was conducted to increase skill and knowledge. The AI intervention was designed to equip L2 graduate students to practice English research writing skills and knowledge. The communication tools were a forum in the virtual classroom and e-mail. AI group intervention consists of 12 weekly two-hour sessions. To align the course contents with the student's needs, the participants were questioned. Also, the high reputation research course was evaluated across several characteristics: the research course Regarding affective acceptance, the first session content is included topics about AI systems, AI application in future life and education, and motivational and practical examples of research writing promotion and facilitation by *Grammarly* application.

Regarding meaningful learning, before each session, the participants were notified about what lessons can be learned from the course. Also, at the end of each session, they were well debriefed about what they have

learned and how they can use this new knowledge or skill in research writing. For skill promotion, the students were asked to do homework and self-assess with App features. To enhance interaction, two social groups were created for each group separately. Beyond the intervention session, useful content and video were be shared through social media. To align course contents with the real world, we designed a social network platform (All Students Are Reviewers: (ASAR)). The participants were requested to share homework in the public group unanimously to receive peer feedback in the single-blind journal format. The research writing syllabus will be developed based on the recent well-established research writing course provided by Elsevier. The intended course was selected according to the course eligibility criteria included: describe the contents in a logical order, combination reliable video, PDF, provide certification through formative assessment, offer authentic skill, provide both linguistic and discipline knowledge. The Grammarly self-assessment skill was instructed, followed by an example to enhance the skill.

**2.4.3.2. NEAI intervention.** The NEAI course was knowledge-based. The NEAI intervention consists of 12 weekly two-hour sessions. The course contents were the same as the AI group, except for *Grammarly* contents. To increase retention, the research team provides the AI course for NEAI participants after the study. The summary of each module content and intervention schedule is demonstrated in Table 2.

**2.4.3.3. Data analysis.** All analyses were conducted with SPSS software version 25 (version 25, SPSS Inc., Chicago, IL), two-tailed with an alpha level of 0.05 to determine statistical significance, following an Intention-to-Treat (ITT) analysis approach. According ITT, the data for all randomized subjects were included in the final report. The last observation-carried-forward method was carried out to handle missing data. An independent *t*-test was generated to investigate whether the students were equivalent at baseline (Time 1). The parametric test of analyzing of covariance (ANCOVA) was used to evaluate the effectiveness of the AI at post-intervention (Time 2). The scores on the baseline were treated as a covariate to control for pre-existing differences between the groups. Levene's test was used to determine normality and homogeneity of variance. Also, the homogeneity of regression slopes assumption was tested. Effect sizes are reported as partial eta squared for ANCOVA. The Between-group comparisons and within-group comparisons were carried out to examine intervention effectiveness between Time 2 and Time 1. Also, Standardized effect size estimates (Cohen's *d*) were calculated for time 1 changes based on means and standard deviations for both groups.

### 3. Results

#### 3.1. Descriptive characteristics at baseline

There were no significant differences in demographic features, age, and other dependent variables at baseline (see Table 1). The AI included 52% women ( $n = 31$ ) and 48% men ( $n = 29$ ). The AI participants' ages ranged from 26 to 38 years ( $M = 31.85$  years,  $SD = 3.55$ ). Also, AI group consists of 28.3% ( $n = 17$ ) in humanity sciences, 28.3% ( $n = 17$ ) technology sciences, and 43.3% ( $n = 26$ ) health sciences. The NEAI included 46.7% females ( $n = 28$ ) and 53.3% males ( $n = 32$ ). NEAI participants' ages ranged from 26 to 39 years ( $M = 32.23$  years,  $SD = 3.16$ ). Also, NEAI group consists of 20 % ( $n = 12$ ) in humanity sciences, 41.7% ( $n = 25$ ) technology sciences, and 38.3% ( $n = 23$ ) health sciences. For learning, mobile devices were preferable to computers. However, for manuscript submissions, the computer (included laptops) were used more. Nine (20%) from the AI group left the experiment before Time 2; 85% ( $n = 51$ ) of the AI group completed the course. Also, 30% ( $n = 18$ ) of the NEAI group dropped out post-intervention. Finally, 30% 77.5% ( $n = 93$ ) of all participants completed the study. On average, participants had a very high degree of adherence, and the program was well tolerated.

**Table 2.** The summary of each module content and intervention schedule.

Session	Content and the Number of Sessions for Module	Homework
One	Writing tools in digital age	
Two	Structuring your article correctly	Write an APA or AMA template
Three	How to prepare your manuscript	
Four	How to write an abstract and improve your article	Write an abstract
Five	Using proper manuscript language	
Six	How to turn your thesis into an article	
Seven	Writing a persuasive cover letter for your manuscript and write an e-mail for collaboration	Write aa cover letter and email for collaboration
Eight	Plagiarism	Find plagiarism in their work
Nine	Report data (APA, AMA), Reference	Write a standard report
Ten	10 tips for writing a truly terrible journal article	
Eleven	full grammarly check <sup>ψ</sup>	Revise an article
Twelve	full grammarly check <sup>ψ</sup>	Revise an article

Note:  $\psi$ : only for intervention group.

### 3.2. Intervention results

The ANCOVA assumptions were examined before submitting the test results. Homogeneity of Variance was tested using Levene's test, indicating insignificance of P-value ( $P > .05$ ). The Homogeneity of Variance assumption was met for all dependent variables. There was no significant interaction between the covariates and the intervention.

**Engagement results:** for Behavioral Engagement, The results revealed a statistically significant main effect for group,  $F(1, 117) = 14.81, p < .001, \eta^2p = .11$ . Between-group, comparisons indicated statistically significant differences between the intervention conditions at post-intervention  $t(118) = 3.87, p < .001$ , and Cohen's  $d = .75, 95\% \text{ CI } [.38, 1.12]$ . For Emotional Engagement, the results revealed a statistically significant main effect for group,  $F(1, 117) = 31.67, p < .001, \eta^2p = .21$ . Between-group, comparisons indicated significant differences between the treatment conditions at post-treatment  $t(118) = 4.50, p < .001$ , and Cohen's  $d = -.82, 95\% \text{ CI } [-1.20, -.45]$ . For Cognitive Engagement, the results revealed a statistically significant main effect for group,  $F(1, 117) = 13.98, p < .001, \eta^2p = .11$ . Between-group comparisons indicated statistically significant differences between the intervention conditions at post-treatment,  $t(118) = 3.84, p < .001$ , and Cohen's  $d = .39, 95\% \text{ CI } [.04, .76]$ .

**Self-Efficacy results:** The results revealed a statistically significant main effect for group,  $F(1, 117) = 20.53, p < .001, \eta^2p = .15$ . Between-group comparisons indicated significant differences between the intervention conditions at post-intervention  $t(118) = 4.79, p < .001$ , and Cohen's  $d = .54, 95\% \text{ CI } [.18, .91]$ .

**Academic emotion results:** for positive academic emotion, The results revealed a statistically significant main effect for group,  $F(1, 117) = 14.26, p < .001, \eta^2p = .10$ . Between-group comparisons indicated statistically significant differences between the intervention conditions at post-intervention  $t(118) = 3.52, p = .003$ , and Cohen's  $d = .44, 95\% \text{ CI } [.08, .80]$ . For Negative academic emotion, The results revealed a statistically significant main effect for group,  $F(1, 117) = 22.83, p < .001,$

$\eta^2p = .16$ . Between-group comparisons indicated statistically significant differences between the intervention conditions at post-intervention  $t(118) = 4.88, p < .001$ , and Cohen's  $d = -.98, 95\% \text{ CI } [-1.36, -.60]$ . Means and standard deviations were calculated at Time 1 and Time 2 (see Table 3).

### 4. Discussion

While most digital writing tools' basic functionality is readily apparent, there may well be preferences or advanced features of which the users may be unaware (Rosell-Aguilar, 2017). This study aimed to examine the efficacy of a group format of the AI application for research writing in L2 Ph.D. students. The results indicated the AI effectiveness in improving self-efficacy, engagement, and academic emotion at post-intervention compared with NEAI. Grammarly can improve students' EFL writing skill (Cavaleri and Dianati, 2016; Parra and Calero, 2019).

Regarding the first hypothesize, the results showed that AI was an effective intervention for improving engagement. Our study supports previous studies on improving writing engagement with AWCF (Koltovskaia, 2020; Zhang, 2020). In line with our findings, technology has a positive influence on multiple indicators of student engagement, such as grit (Schindler et al., 2017). Grammarly provides immediate feedback and revision. Immediate feedback positively predicts engagement in web-based courses. The revision may motivate the students to revise by receiving the technology scores (Moore and MacArthur, 2016). For research writing, the findings indicated no statistically significant difference in the engagement scores for non-equipped L2 students under the same instructional procedure.

The engagement is a crucial factor for a technology pedagogical evaluation. This real-time formative feedback, followed by practical and instructional examples, provides new possibilities for learners' more personalized experiences. There is evidence that instructional feedbacks may enhance writing problem-solving (Li et al., 2014) and self-regulatory

**Table 3.** Mean and Standard deviation at pre-treatment and post-treatment.

Item	NEAI		AI	
	Time 1	Time 2	Time 1	Time 2
	M (SD)	M (SD)	M (SD)	M (SD)
Self-efficacy	29.63 (4.78)	30.97 (4.58)	31.71 (5.60)	34.48 (6.23)
Engagement	13.63 (5.12)	14.28 (5.64)	13.71 (5.23)	17.85 (4.39)
Positive emotion	2.94 (.59)	3.00 (.87)	2.88 (.73)	3.35 (.70)
Negative emotion	2.98 (.42)	2.89 (.58)	2.85 (.64)	2.25 (.72)

Note: M: Mean, SD: Standard Deviation, Time 1:pre-treatment, Time 2: post-treatment.

strategies (Wang et al., 2013). Self-evaluation alludes to the tendency and capability to precisely assess and monitor student's knowledge while learning. Self-Evaluation and revision review are associated with writing proficiency and higher writing scores. However, the level of Self-assessment skills may lead to select students' optimal or sub-optimal learning strategies. Research indicates that students' self-assessment sometimes is associated with overestimation. These overestimated and erroneous assessments can lead to suboptimal strategy selection, while an accurate self-assessment is related to productive learning strategies and social, behavioral engagement such as help-seeking. The type of feedback also impacts the students' cognitive load. The kind of assessment affects students' cognitive load. Automated scoring can lead to less cognitive load when assessment focuses on improve fluency and increase control of the text production processes (Kellogg and Raulerson, 2007).

The results showed that AI was an effective intervention for improving individual acceptance factors regarding the second hypothesis. The results showed that AI was an effective intervention for enhancing self-efficacy and academic emotions in L2 students. Our findings are consistent with previous research on AI's efficacy in L2 students (Parra and Calero, 2019; Cotos, 2011). In English writing, learners with higher self-efficacy levels are more likely to put more effort, exert more persistence, and thus have higher writing outcomes (Usher and Pajares, 2008). Using AWE can improve the students' confidence in writing, specifically when they receive positive feedback. Intelligent feedback can reinforce writing autonomy when allowing the student to inspect their errors, identify the incorrect writing patterns, and reformulate the errors, mainly when no human support is available. Hegelheimer and Lee (2013) report positive preliminary findings of both students' and teachers' perspectives. According to Teachers' reports, their students explained that AWE was useful to understand their weaknesses better. According to students' self-report, using AWE, they were better at detecting and correct writing errors than before.

The results showed that AI was an effective intervention for improving positive and decreasing negative emotions in L2 students (Su et al., 2019). AIs provide in-depth analytical assessments and immediate feedback, which can induce a series of emotions in academic settings. Mainly, self-assessment, followed by self-referential feedback, can generate positive achievement emotions (i.e., hope and pride; Vogl and Pekrun, 2015). The normative Assessments administered by teachers can induce negative emotions (e.g., anxiety). Grammarly Provides the opportunity to self-correct the tasks before a summative assessment via revision. It helps to highlight areas where the students are going wrong before they get graded, meaning students have the opportunity to improve their work. Self-regulatory skills are critical to managing negative emotions associated with writing (Wei and Chou, 2020). There is evidence that giving more opportunities (e.g., chance a re-test) can diminish test anxiety (Zeidner, 2014). When negative emotions create a pessimistic perceptual attitude (Pekrun et al., 2002), they divert the learner's attention to factors irrelevant to the task, which activate intrusive thoughts priority to a concern for well-being rather than for learning. Grammarly's impact on negative academic emotions related to assessment (anxiety) can be considered as this AI's affect students' well-being (Zeidner, 2014).

#### 4.1. Pedagogical implication

From a pedagogical perspective, the benefits of using ICT in teaching can be learning effectiveness, satisfaction, and efficiency. An accurate self-assessment is associated with productive learning strategies and social, behavioral engagement, such as help-seeking. They develop self-regulation and self-reflection abilities, become assessment capable, and have their motivation and confidence boosted as a result. They also take responsibility for learning and become autonomous learners.

According to the pedagogical approach related to sustainability, current assessment in higher education is inadequate to prepare students for a lifetime of learning (Boud, 2000). Higher education institutions

must reconsider their assessment methods to equip their learners with new skills and competencies for sustainable assessment. This assessment feature helps students become more active learners to manage to learn and assess the works (Boud and Soler, 2016). At the same time, the institutions move away from teacher-centered learning to student-centered. In traditional L2 writing environments, the teacher role is predominant, which generates dissatisfaction, anxiety, and boredom. To shift from teacher-centered to student-learner-centered, feedback and assessment can play a valuable role. Technology can provide an opportunity for teachers to reduce the workload with student self-assessment. Self-assessment can develop students' motivation, increase their ability to lead their learning (McMillan and Hearn, 2008). There is evidence that an online grammar checker is useful for low-proficient L2 learners' writing (Grimes and Warschauer, 2010). Compared to teacher feedback, AI's significantly reduced writing errors (Vajjala, 2018).

Research indicates that tech-powered digital tools can support students to become self-directed learners. These technologies can enhance the creation of knowledge and develop new competencies (Scardamalia and Bereiter, 2015). For research writing, while self-efficacy is critical for all portions of writing and also for technology integration. The findings of the current study are crucial. Grammarly can be utilized to enhance self-efficacy. For students with a fixed mindset, Feedbacks can modify the students' views of intelligence. When students see ability and skill can be developed through practice, (b) practice is central to gain the skill, and (c) mistakes are inherently part of the learning process (Shute, 2008).

Emotions are an integral part of learning success can potentially be led by teachers and other educators who will have developed pedagogical understanding to know how to push young people without ridicule or demotivation. To enhance instructors' effectiveness of learning, instructors are encouraged to identify affective-stimulating (e.g., negative emotions) and provide a supportive learning environment so that they can devote their complete working memory resources to the learning tasks (Chen and Chang, 2017).

For mobile learning, our findings suggest that technology can be positively influenced by technology in line with recent research (e.g., Wilson and Roscoe, 2020; Bernacki et al., 2020). Within the pedagogical perspective, technology can reformulate the learning environment. The media are commonplace for learners outside school. Mobile technologies can provide teachers with flexible teaching opportunities and ongoing formative assessment (Holstein et al., 2017; Reeves et al., 2017), specifically through social media. Social media can be utilized as a sharing information platform and creative learning environment. In the present study, social media considers a single-blind peer-review platform. Through the platform, students can promote relevant skills about the assessments and participate in the collaborative learning experience.

#### 5. Limitation

The results from this trial should be interpreted in the context of several limitations. The findings revealed low to moderate statistically significant differences for the AI group. Also, the results showed no statistically significant differences in the NEAI group. Also, the students are usually highly-engaged, self-motivated, and gritty students. So, there is not surprising that the traditional course can change no statistically significant differences. Next, data were based on self-report, and thus suffer the limitations associated with all self-reported data. As previously outlined, the participants were generally highly-engaged students, which can be enhanced their abilities to gain more of the course and diminish the results' generalizability.

In the present study, the main AI application limitation is that AI learns from given data. There is no other way that knowledge can be integrated, unlike human learning. Therefore, any inaccuracies in the data will be reflected in the results. Pedagogically, there is currently lack of the supports to teachers in the integration of AI technologies. Also, AI competencies and standard literacy must be determined, specifically. This contribution is critical. Because; AI application need to an optimum

knowledge and skills in computer literacy, ICT literacy, information literacy, and media literacy. In future studies, there would be value in adding qualitative approach into future trials to establish the AI Applications' acceptability for both teachers and students.

## 6. Conclusion

This study's major strength is the use of randomized controlled trials and prior-powered analyses with a control design. Also, baseline data in the format and analysis of a randomized trial will be considered using a real-world example. This is a more robust design than the simple 'after only' design, and often a substantially more powerful one. Overall, our findings the utility of applying an AI-powered writing tool to improve self-efficacy, engagement and emotion, emotions in the EFL context. Results suggest that *Grammarly* may affect cognitive, non-cognitive, and emotional domains of learning. The results indicate that AI application could be an efficient assist-writing tool in English academic writing for non-native postgraduate students.

## Declarations

### Author contribution statement

Nabi Nazari: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Muhammad Salman Shabbir: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Roy Setiawan: Contributed reagents, materials, analysis tools or data; Wrote the paper.

### Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Data availability statement

Data will be made available on request.

### Declaration of interests statement

The authors declare no conflict of interest.

### Additional information

No additional information is available for this paper.

## References

- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84 (2), 191.
- Bandura, A., 2015. On deconstructing commentaries regarding alternative theories of self-regulation. *J. Manag.* 41 (4), 1025–1044.
- Bernacki, M.L., Greene, J.A., Crompton, H., 2020. Mobile technology, learning, and achievement: advances in understanding and measuring the role of mobile technology in education. *Contemp. Educ. Psychol.* 60. November 2019.
- Boud, D., 2000. Sustainable assessment for long term learning. *Stud. Cont. Educ.* 22 (2), 151–167.
- Boud, D., Soler, R., 2016. Sustainable assessment revisited. *Assess Eval. High Educ.* 41 (3), 400–413.
- Brinkerhoff, J., 2006. Effects of a long-duration, professional development academy on technology skills, computer self-efficacy, and technology integration beliefs and practices. *J. Res. Technol. Educ.* 39 (1), 22–43.
- Bruning, R., Dempsey, M., Kauffman, D.F., McKim, C., Zumbrunn, S., 2013. Examining dimensions of self-efficacy for writing. *J. Educ. Psychol.* 105 (1), 25–38.
- Campbell, M., 2019. Teaching academic writing in higher education. *Educ. Q. Rev.* 2 (3), 608–614.
- Cavaleri, M., Dianati, S., 2016. You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *J. Acad. Lang. Learn.* 10 (1), A223–A236.
- Chen, I.-J., Chang, C.-C., 2017. Cognitive load theory: an empirical study of anxiety and task performance in language learning. *Electron. J. Res. Educ. Psychol.* 7 (18), 729–746.
- Cheung, S.K.S., 2015. A case study on the students' attitude and acceptance of mobile learning. *Commun. Comput. Informat. Sci.* 45–54.
- Christensen, R., Knezek, G., 2017. Validating the technology proficiency self-assessment for 21st century learning (TPSA C21) instrument. *J. Digit. Learn. Teacher Educat.* 33 (1), 20–31.
- Cotos, E., 2011. Potential of automated writing evaluation feedback. *CALICO J.* 28 (2), 420–459.
- Dumbauld, J., Black, M., Depp, C.A., Daly, R., Curran, M.A., Winegarden, B., Jeste, D.V., 2014. Association of learning styles with research self-efficacy: study of short-term research training program for medical students. *CTS-Clin. Translat. Sci.* 7 (6), 489–492.
- Fadhilah, U., Julia, H., Saribu, D., 2019. Effectiveness of grammarly application for writing English Abstract, 8 (12), 163–166.
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2009. Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41 (4), 1149.
- Fook, 2010. Authentic assessment and pedagogical strategies in higher education. *J. Soc. Sci.* 6 (2), 153–161.
- Fredricks, J.A., Blumenfeld, P.C., Paris, A.H., 2004. School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74 (1), 59–109.
- Fredricks, J.A., Blumenfeld, P., Friedel, J., Paris, A., 2005. School Engagement. In: Moore, K.A., Lippman, L.H. (Eds.), *The Search Institute series on developmentally attentive community and society. What do children need to flourish: Conceptualizing and measuring indicators of positive development.* Springer Science + Business Media, pp. 305–321.
- Fredricks, J.A., McColskey, W., 2012. The measurement of student engagement: a comparative analysis of various methods and student self-report instruments. In: Christenson, S.L., et al. (Eds.), *Handbook of Research on Student Engagement.* Springer Science Business, pp. 763–782.
- Graesser, A.C., D'Mello, S.K., Strain, A.C., 2015. Emotions in advanced learning technologies. *Int. Handbook Emot. Educat.*
- Grimes, D., Warschauer, M., 2010. Utility in a fallible tool: A multi-site case study of automated writing evaluation. *J. Technol. Learn. Assessment* 8 (6).
- Hanauer, D.L., Sheridan, C.L., Englander, K., 2019. Linguistic injustice in the writing of research articles in English as a second language : data from Taiwanese and Mexican Researchers.
- Hegelheimer, V., Lee, J., 2013. The role of technology in teaching and researching writing. In: Thomas, M., Reinders, H., Warschauer, M. (Eds.), *Contemporary Computer-Assisted Language Learning.* Bloomsbury Academic, pp. 287–302.
- Holstein, K., McLaren, B.M., Aleven, V., 2017. Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference.* ACM, pp. 257–266.
- Hong, A.J., Kim, H.J., 2018. College students' digital readiness for academic engagement (DRAE) scale: scale development and validation. *Asia-Pac. Educat. Researcher* 27 (4), 303–312.
- Jonassen, D.H., Howland, J., Moore, J., Marra, R.M., 2003. *Learning to Solve Problems with Technology: A Constructivist Perspective*, 2nd. Ed. Merrill/PrenticeHall, Columbus, OH.
- Kellogg, R.T., Raulerson, B.A., 2007. Improving the writing skills of college students. *Psycho. Bull. Rev.* 14 (2), 237–242.
- Koltovskaia, S., 2020. Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: a multiple case study. *Assessing Writing* 44 (April 2020), 100450.
- Kopp, B., Schulze, E., Mandl, H., 2008. Which Design Principles Influence Acceptance and Motivation in Professional E-Learning? beyond Knowledge: the Legacy of Competence, pp. 83–92.
- Li, Z., Link, S., Ma, H., Yang, H., Hegelheimer, V., 2014. The role of automated writing evaluation holistic scores in the ESL classroom. *System* 44 (1), 66–78.
- Lim, F.V., Phua, J., 2019. Teaching writing with language feedback technology. *Comput. Compos.* 54, 102518.
- Lishinski, A., Yadav, A., Good, J., Embody, R., 2016. Learning to program: gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In: *Proceedings of International Computing Education Research (ICER '16)*, Melbourne, Australia, pp. 211–220.
- Lobue, V., Pérez-Edgar, K., Buss, K.A., 2019. *Handbook of emotional development.* Cham. Springer, Switzerland.
- Mandl, H., Bürg, O., 2008. Acceptance of e-learning at the workplace; the influence of institutional framework. In: Nijhof, W.J., Nieuwenhuis, L.F.M. (Eds.), *The Learning Potential of the Workplace.* Sense Publishers, pp. 213–224, 2008.
- McMillan, J.H., Hearn, J., 2008. Student self-assessment: the key to stronger student motivation and higher achievement, educational. *HORIZONS* 87, 40–49.
- McRobert, C.J., Hill, J.C., Smale, T., Hay, E.M., van der Windt, D.A., 2018. A multi-modal recruitment strategy using social media and internet-mediated methods to recruit a multidisciplinary, international sample of clinicians to an online research study. *Plos One* 13 (7), e0200184.
- Molin, F., Haelermans, C., Cabus, S., Groot, W., 2020. The effect of feedback on metacognition - a randomized experiment using polling technology. *Comput. Educ.* 152 (October 2019), 103885.



- Moore, N.S., MacArthur, C.A., 2016. Student use of automated essay evaluation technology during revision. *J. Writ. Res.* 8 (1), 149–175.
- Nazari, N., Far, D.M., 2019. The relationship between teaching skills, academic emotion, academic stress and mindset in university student academic achievement prediction: a PLS-SEM approach. *J. Intellect. Disabil. - Diagn. Treat.* 7 (3), 119–133.
- Nazari, N., Griffiths, M.D., 2020. Psychometric validation of the Persian version of the emotional style questionnaire. *Curr. Psychol.* 1–13. Advance online publication.
- Nazari, N., 2020. Developing and validating the teacher self-efficacy for teaching students with autism spectrum disorder (TSE-ASD) scale. *J. Intellect. Disabil. - Diagn. Treat.* 8 (4), 673–678.
- Pajares, F., Schunk, D.H., 2002. Self and self-belief in psychology and education: A historical perspective. In: Aronson, J. (Ed.), *Improving Academic Achievement: Impact of Psychological Factors on Education*. Academic Press, pp. 3–21.
- Parra, G.L., Calero, S.X., 2019. Automated writing evaluation tools in the improvement of the writing skill. *Int. J. InStruct.* 12 (2), 209–226.
- Pedro, L.F.M.G., de O. Barbosa, C.M.M., das N. Santos, C.M., 2018. A critical review of mobile learning integration in formal educational contexts. *Int. J. Educat. Technol. Higher Educat.* 15 (1).
- Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., Perry, R.P., 2011. Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemp. Educ. Psychol.* 36 (1), 36–48.
- Pekrun, R., Goetz, T., Titz, W., Perry, R.P., 2002. Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ. Psychol.* 37 (2), 91–105.
- Pekrun, R., Elliot, A.J., Maier, M.A., 2009. Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *J. Educ. Psychol.* 101 (1), 115–135.
- Pintrich, P.R., 2000. The role of goal orientation in self-regulated learning. In: Boekaerts, M., Pintrich, P.R., Zeidner, M. (Eds.), *Handbook of Self-Regulation*. Academic Press, pp. 451–502.
- Popenici, S.A.D., Kerr, S., 2017. Exploring the impact of artificial intelligence on teaching and learning in higher education. *Res. Pract. Technol. Enhanc. Learn. (RPTEL)* 12 (1).
- Rahimi, M., Zhang, L.J., 2018. Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Read. Writ.* 32 (3), 761–786.
- Reeves, J.L., Gunter, G.A., Lacey, C., 2017. Mobile learning in pre-kindergarten: Using student feedback to inform practice. *J. Educ. Technol. Soc.* 20 (1), 37–44.
- Rosell-Aguilar, F., 2017. State of the app: a taxonomy and framework for evaluating language learning mobile applications. *CALICO J.* 34 (2), 243–258.
- Ryan, R.M., Deci, E.L., 2018. *Self-determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford, New York.
- Sadiq, J.M., 2017. Anxiety in English language learning: a case study of English language learners in Saudi Arabia. *Engl. Lang. Teach.* 10 (7), 1.
- Scardamalia, M., Bereiter, C., 2015. Education in an open informational world. In: Scott, R., Kosslyn, S. (Eds.), *Emerging Trends in the Social and Behavioural Sciences: An Interdisciplinary, Searchable, and Linkable Resource*. Wiley, Hoboken, pp. 1–15.
- Schindler, L.A., Burkholder, G.J., Morad, O.A., Marsh, C., 2017. Computer-based technology and student engagement: a critical review of the literature. *Int. J. Educat. Technol. Higher Educat.* 14 (1).
- Shute, V.J., 2008. Focus on formative feedback. *Rev. Educ. Res.* 78 (1), 153–189.
- Skinner, E., Pitzer, J., Brule, H., 2014. The role of emotion in engagement, coping, and the development of motivational resilience. *Int. Handbook Emot. Educat.* 331–347.
- Su, Z., Miao, L., Man, J., 2019. Artificial intelligence promotes the evolution of English writing evaluation model. *IOP Conf. Ser. Mater. Sci. Eng.* 646 (1).
- Taguma, M., Senior Analyst, M., Feron, E., Meow Hwee, L.I.M., 2018. Future of education and skills 2030: conceptual learning framework. November.
- Usher, E.L., Pajares, F., 2008. Self-efficacy for self-regulated learning: A validation study. *Educ. Psychol. Measure.* 68 (3), 443–463.
- Vajjala, S., 2018. Automated assessment of non-native learner essays: investigating the role of linguistic features. *Int. J. Artif. Intelligence Educ.* 28 (1), 79–105.
- VogL, E., Pekrun, R., 2015. Emotions that matter to achievement: student feelings about assessment. In: (GavinBrown, T.L., LoisHarris, R. (Eds.), *Handbook of Human and Social Conditions in Assessment*, 2016. Taylor & Francis, pp. 111–129.
- Voogt, J., Knezek, G., Christensen, R., Lai, K.-W., 2018. Developing an understanding of the impact of digital technologies on teaching and learning in an ever-changing landscape. In: *Springer International Handbooks of Education*, vol. 1, pp. 3–12.
- Wang, C., Bai, B., 2017. Validating the instruments to measure ESL/EFL learners' Self-efficacy beliefs and self-regulated learning strategies. *Tesol Q.* 51, 931–947.
- Wang, Y.-J., Shang, H.-F., Briody, P., 2013. Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Comput. Assist. Lang. Learn.* 26 (3), 234–257.
- Wang, Z., Gong, S., Xu, S., Hu, X., 2019. Computers & Education Elaborated feedback and learning: examining cognitive and motivational influences. *Comput. Educ.* 136 (April), 130–140.
- Wei, H.C., Chou, C., 2020. Online learning performance and satisfaction: do perceptions and readiness matter? *Dist. Educ.* 41 (1), 48–69.
- Wible, D., Kuo, C.-H., Chien, F., Liu, A., Tsao, N.-L., 2001. A Web-based EFL writing environment: integrating information for learners, teachers, and researchers. *Comput. Educ.* 37 (3–4), 297–315.
- Wilson, J., Roscoe, R.D., 2020. Automated writing evaluation and feedback: multiple metrics of efficacy. *J. Educ. Comput. Res.* 58 (1), 87–125.
- Wosnitza, M., Volet, S., 2005. Origin, direction and impact of emotions in social online learning. *Learn. InStruct.* 15 (5), 449–464.
- Yu, S., Zhou, N., Zheng, Y., Zhang, L., Cao, H., Li, X., 2019. Evaluating student motivation and engagement in the Chinese EFL writing context. *Stud. Educ. Eval.* 62 (19), 129–141.
- Zawacki-Richter, O., Marín, V.I., Bond, M., Gouverneur, F., 2019. Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int. J. Educat. Technol. Higher Educat.* 16 (1).
- Zeidner, M., 2014. Anxiety in education. In: Pekrun, R., Linnenbrink, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York, pp. 265–288.
- Zimmerman, B. J., & Labuhn, A. S. (2012). Self-regulation of learning: Process approaches to personal development. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA handbooks in psychology®. APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (p. 399–425). American Psychological Association.
- Zhang, Z., 2020. Engaging with automated writing evaluation (AWE) feedback on L2 writing: student perceptions and revisions. *Assess. Writ.* 43 (November), 100439.
- Zimmerman, B., Kitsantas, A., 2007. Reliability and validity of self-efficacy for learning form (SELF) scores of college students. *J. Psychol.* 215 (3), 157–163.