

The Role of Prompt Engineering in Enhancing Critical Thinking Disposition and Engagement in Face-to-Face Group Discussions

Tala Bagheri Malek Abadi ¹ , Goodarz Shakibaei ^{2*} 

¹ M.A Student, English Department, Ahv.C. Islamic Azad University, Ahvaz, Iran

² Assistant Professor, Department of English Language Teaching, Ahv.C. Islamic Azad University, Ahvaz, Iran



10.22080/iselt.2026.32546.1219

Received

July 27, 2026

Accepted

September 28, 2026

Available online

October 5, 2026

Keywords:

Critical thinking
Disposition, EFL
Learners, Face-To-Face
Group Discussions,
Prompt Engineering,
Student Engagement.

Abstract

Prompt engineering is increasingly used to structure learner responses, yet its pedagogical value in face-to-face English as a foreign language (EFL) discussion remains underexplored. This study examined whether systematically designed instructor prompts could improve Iranian EFL learners' engagement and critical thinking disposition (CTD) during classroom group discussions. Fifty intermediate-level female learners were randomly assigned to a prompt-supported discussion group or a conventional discussion group and completed pretest and posttest measures before and after an 8-week intervention. Paired-samples t-tests indicated significant within-group improvement, whereas independent-samples t-tests showed significantly higher posttest engagement and CTD in the prompt-supported group. The findings are consistent with such prompts helping to organize participation and create conditions conducive to analytical and reflective engagement. By extending prompt engineering beyond predominantly human-AI interaction, the findings suggest that structured prompt design may serve as a pedagogically grounded resource for face-to-face EFL discussions that are both cognitively substantive and participatory.

1. INTRODUCTION

Prompt engineering refers to the deliberate formulation and refinement of prompts to elicit responses aligned with particular goals. Although the concept is now closely associated with generative artificial intelligence (GenAI) and large language models (LLMs), its underlying principle is pedagogically relevant beyond human-AI interaction because the structure, specificity, and cognitive demands of a prompt can shape the nature of the response it elicits (Ding et al., 2022; Knoth et al., 2024; Liu et al., 2023). Design-oriented research likewise shows that prompt formulation can meaningfully influence generated outputs (Liu & Chilton, 2022). Recent educational studies further indicate that the value of GenAI depends partly on how interaction is

* **Corresponding Author:** Goodarz Shakibaei, Assistant Professor, Department of English Language Teaching, Ahv.C. Islamic Azad University, Ahvaz, Iran, **Email:** goodarzshakibaei@gmail.com



pedagogically structured, including the relationship between AI-mediated support and peer dialogue (Bozorgian & Rahimi, 2025; Deguchi et al., 2026). Here, however, prompt engineering is operationalized specifically as the purposeful design and sequencing of instructor-generated prompts for face-to-face English as a foreign language (EFL) group discussion. These prompts are intended to encourage learners to explain, compare, justify, evaluate, and reconsider ideas rather than simply recall information.

This approach is particularly relevant to classroom discussion because the quality of a question can influence both learners' participation and the depth of their reasoning. Prompts that request factual information may invite brief responses, whereas prompts requiring justification, comparison, evaluation, or consideration of alternative perspectives can encourage more elaborated responses. Research in English-language classrooms indicates that higher-level questioning can promote participation and critical dialogue when learners receive appropriate instructional support (DeWaele, 2015). Accordingly, the pedagogical value of prompt engineering lies not in increasing the number of questions, but in designing prompts that provide sufficient interactional direction while stimulating cognitively substantive discussion.

Engagement is a critical factor influencing the quality of students' learning experiences and academic success. It is commonly conceptualized as a multidimensional construct comprising behavioral, cognitive, and affective components (Fredricks et al., 2004). Behavioral engagement concerns observable participation and effort (Kahu, 2013), cognitive engagement involves the mental investment required to understand and respond meaningfully to learning content (Sang & Hiver, 2021), and emotional engagement reflects affective responses such as interest and enjoyment (Erdemutu et al., 2024). Although digitally-mediated research has also examined sustained participation, cognitive investment, and affective responses (Jones et al., 2022; Ni et al., 2020), such evidence is contextually different from face-to-face classroom discussion and is used here only to illustrate the multidimensional nature of engagement. Engagement has been associated with persistence and academic achievement (Kahu & Nelson, 2018), while language-learning research likewise links behavioral, cognitive, and affective engagement with sustained involvement and motivation (Ebadi et al., 2024). Related evidence from Iranian EFL settings suggests that broader classroom motivational conditions can shape learners' willingness to participate and attend (Rahimi et al., 2026). Here, engagement, therefore, refers specifically to learners' behavioral participation, cognitive investment, and affective involvement during face-to-face group discussion.

The cognitive dimension of engagement is closely related to critical thinking (CT), which involves evaluating evidence, considering alternative perspectives, questioning assumptions, and reaching reasoned judgments rather than accepting information uncritically (Geever et al., 2021; Pescatore, 2007; Zou'bi, 2021). Educational research indicates that CT benefits from explicit instructional support and from learning environments that require learners to articulate reasons, evaluate information, and reconsider conclusions (Polat & Aydin, 2020). Group activities, problem-solving, source evaluation, and questioning can provide such opportunities (Aizikovitsh-Udi & Cheng, 2015; Florea & Hurjui, 2015; Hayes & Devitt, 2008). CT has also been considered within broader life-skills frameworks (Erawan, 2010), while critical thinking disposition has been examined in EFL contexts (Hoorijani & Heidari Tabrizi, 2024). In the present study, the critical-thinking outcome is operationalized as critical thinking disposition (CTD), specifically learners' openness to alternative perspectives and reflective scepticism. This distinction is important because the study examines learners' disposition toward critical engagement rather than performance on a separate CT skills test. Although some of the cited studies predate the current GenAI-oriented use of the term prompt engineering, they remain relevant to the pedagogical mechanism of structured questioning that underpins the intervention.

Structured prompts may connect engagement and CT through several complementary processes. Prompts that require learners to explain or justify a response encourage them to articulate the reasoning underlying their position. Questions that introduce alternative viewpoints invite reconsideration of initial assumptions. Prompts requiring learners to challenge, extend, or respond to peers can also foster reciprocal interaction. Problem-oriented prompts may further encourage learners to compare solutions, evaluate evidence, and negotiate conclusions. In this way, prompt design offers a means of promoting active participation and analytical reasoning simultaneously (DeWaeltsche, 2015).

These processes are particularly important in EFL group discussions, where learners must manage language production alongside reasoning, turn-taking, and responsiveness to others. A general instruction to discuss a topic may provide insufficient direction for learners who are uncertain about how to develop their contributions. Conversely, cognitively demanding questions without adequate support may result in limited participation. Systematically designed prompts offer a clearer pathway by specifying cognitive operations such as explaining, comparing, evaluating, challenging, or reconsidering ideas without prescribing what learners should conclude.

Despite growing interest in prompt engineering, contemporary research has focused predominantly on interaction with GenAI and LLMs, including AI literacy, prompt-writing strategies, prompting frameworks, and the quality of human-AI interaction (Knoth et al., 2024; Lee & Palmer, 2025). Within EFL research, Hwang et al. (2025) similarly showed that the clarity, contextualization, and strategic construction of learner-generated prompts influence the quality of interaction with ChatGPT. A related classroom literature has established the value of higher-level questioning, collaborative activity, problem-solving, and evaluation for engagement and CT-oriented learning (Aizikovitsh-Udi & Cheng, 2015; DeWaeltsche, 2015; Florea & Hurjui, 2015; Hayes & Devitt, 2008). What remains less clear is whether the logic of systematic prompt design can be transferred from predominantly human-AI interaction to face-to-face EFL peer discussion in a way that supports both engagement and CTD.

Examining these outcomes together is important because visible participation does not necessarily entail reflective or evaluative thinking, whereas cognitively demanding tasks do not automatically sustain participation. Accordingly, this study examines whether instructor-designed prompts embedded in face-to-face classroom discussion are associated with stronger engagement and CTD than conventional group discussion. This gap is examined further in the empirical review.

2. RESEARCH BACKGROUND

Theoretical Background

This study draws on complementary constructivist and sociocultural perspectives, self-determination theory (SDT), social cognitive theory (SCT), and cognitive load theory (CLT) to explain how systematically designed prompts may support EFL learners' engagement and CTD during face-to-face group discussions. From a constructivist perspective, learners actively develop understanding by connecting new information with prior knowledge and reorganizing their interpretations through experience and reflection (Piaget, 1973). Vygotsky's (1978) sociocultural perspective adds an explicitly social dimension by emphasizing that higher cognitive processes develop through mediated interaction with others. These principles are pertinent to structured classroom prompting because prompts can direct learners toward higher-order reasoning without determining the substance of their responses. When prompts progress from relatively accessible questions to more demanding analytical tasks, they can provide temporary instructional support within learners' zone of proximal development (ZPD). Peer discussion, then, allows learners to

articulate their reasoning, encounter alternative perspectives, and refine their understanding through interaction (Tavassoli et al., 2024). Thus, structured prompting is theoretically relevant to CT not because prompts supply answers, but because they organize opportunities for learners to construct, examine, and revise meaning collaboratively (Piaget, 1973; Vygotsky, 1978).

SDT provides a complementary explanation for engagement. Deci and Ryan (1985) propose that the quality of learners' motivation is closely related to the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. Subsequent SDT scholarship has demonstrated the educational importance of learning environments that support these needs (Ryan & Deci, 2020). Within group discussion, systematically designed prompts may support autonomy when learners are free to formulate and defend their own positions rather than reproduce predetermined responses. Competence can be supported when prompts clarify the intellectual expectations of the task and provide an attainable pathway from initial participation toward more elaborated contributions. Relatedness may be supported when prompts encourage learners to attend to, respond to, extend, or constructively question their peers' ideas. These processes offer a plausible basis for behavioral, cognitive, and emotional engagement because learners are given opportunities to participate, invest effort in reasoning, and remain socially involved in the developing discussion. Importantly, such guidance should remain autonomy-supportive rather than controlling. The theoretical value of structured prompting, therefore, lies in providing sufficient direction to facilitate participation and preserve learners' ownership of the ideas they contribute (Deci & Ryan, 1985; Ryan & Deci, 2020).

The social processes underlying prompt-supported discussion can be further understood through Bandura's (1986) SCT. Within this framework, learning emerges from reciprocal interactions among personal factors, behavior, and the social environment, and observation and modeling provide important routes through which individuals acquire and adapt new patterns of behavior. During group discussions, learners can observe how peers interpret questions, justify claims, respond to disagreement, and reformulate arguments. Structured prompts can make these reasoning practices more visible by establishing shared expectations for the kinds of contributions required. For instance, a prompt requesting evidence for a position exposes learners not only to different viewpoints, but also to different ways of supporting those viewpoints. As learners listen to and respond to one another, peer contributions can become resources for comparison, self-monitoring, and revision. This reciprocal process is relevant to both outcomes examined here. Engagement may be strengthened as learners participate in and respond to an evolving social exchange, whereas critical engagement is exercised as they examine competing interpretations, evaluate reasons, and modify their positions when appropriate. From this perspective, prompt design constitutes one element of the classroom environment shaping interaction, although the direction and quality of the discussion continue to depend on learners' active responses and reciprocal exchanges (Bandura, 1986).

Finally, CLT explains why the clarity and sequencing of prompts may matter when EFL learners must simultaneously interpret content, formulate ideas in English, follow peers' contributions, manage interaction, and engage in reasoning. Sweller (1988) emphasizes the limited capacity of working memory and the importance of instructional design in preventing cognitive resources from being consumed by demands that do not contribute to learning. Later developments likewise stress the distinction between task-inherent demands and avoidable extraneous processing (Sweller et al., 2019). In face-to-face discussion, vague or poorly structured tasks may impose unnecessary demands because learners must determine both what to discuss and how to organize their contributions. Well-designed prompts can reduce such ambiguity by directing attention toward relevant reasoning operations without removing the intellectual challenge inherent in the task. Taken together, constructivist and sociocultural perspectives explain how mediated

interaction can support the construction and revision of understanding, SDT clarifies how autonomy, competence, and relatedness can sustain engagement, SCT accounts for modeling and reciprocal peer influence, and CLT explains how instructional structure can help direct limited cognitive resources toward task-relevant processing (Sweller, 1988; Sweller et al., 2019).

Empirical Background

Recent research in language education indicates that prompt design and instructional framing can influence AI-mediated learning interactions, although most evidence concerns learner-AI exchanges. In directly relevant EFL work, Hwang et al. (2025) found that clearer and more interconnected prompting strategies were associated with greater satisfaction with ChatGPT responses, whereas Luo et al. (2026) showed that increasingly sophisticated prompts improved the accuracy and coverage of ChatGPT-generated corrective feedback in second language (L2) writing. Other studies outside face-to-face EFL discussion likewise show that explicit prompt instruction can strengthen prompt-writing competence, AI self-efficacy, confidence, autonomous learning, or output quality (Gastineau, 2024; Grabb, 2023; Jacobsen & Weber, 2025; Mzwri & Turcsányi-Szabó, 2025; Woo et al., 2026). In EFL writing, Bozorgian and Rahimi (2025) found complementary cognitive and interpersonal functions for peer and ChatGPT-4o feedback, and Deguchi et al. (2026) showed that higher-order engagement with GenAI can depend on deliberate instructional design and peer dialogue. These studies underscore the broader importance of prompt and interaction design, but they do not directly test systematically engineered prompts in face-to-face EFL peer discussion.

Research on questioning and collaborative language learning provides more closely related evidence for the two outcomes examined here. DeWaeltsche (2015) showed that higher-level questioning in university English courses could encourage student engagement and more critical classroom dialogue. Wu et al. (2026) similarly found that chatbot-supported collaborative learning enhanced engagement during speaking-oriented EFL instruction. With respect to CT, Liang and Wu (2024) reported that learners perceived a structured ChatGPT-supported program as cognitively useful, although pretest-posttest analyses did not reveal a significant improvement in CT scores; by contrast, Shang et al. (2025) reported increasingly sophisticated questioning patterns in a ChatGPT-supported competitive format. Aydın Yıldız (2023) reported motivational benefits in a ChatGPT-supported EFL activity, whereas Schmidt-Fajlik (2023) examined ChatGPT as a grammar-checking resource for Japanese EFL learners. These studies serve only as peripheral contextual evidence rather than direct support for engagement or CTD. Overall, the evidence suggests that the effects of structured interaction depend substantially on how activities are pedagogically organized.

Evidence from Iran further underscores the relevance of explicit CT-oriented instruction and highlights a more specific unresolved issue. Bakhtiari Moghadam et al. (2023) found that a structured CT intervention improved Iranian EFL learners' CT, reading comprehension, and perceptions of classroom climate. At a broader level, Khalili et al.'s (2024) meta-analysis reported an overall association between CT and L2 learning in the Persian EFL context, although substantial heterogeneity and potential publication bias called for cautious interpretation and more focused research. Most recently, Nasseh et al. (2026) showed through classroom-based action research that inquiry-oriented EFL instruction in Iran can support reasoned judgment and critical participation through collaborative classroom practices. These studies suggest that CT is responsive to purposeful pedagogy in Iranian EFL settings (Farsinejad & Haddad Narafshan, 2024), but they do not examine systematically engineered discussion prompts as an intervention or determine whether such prompts can improve engagement and CT concurrently.

The available evidence leaves a specific empirical gap: previous work has not established whether systematically engineered prompts embedded in face-to-face EFL peer discussion can jointly support classroom engagement and CTD. To address this unresolved intersection, the study compares prompt-supported and conventional discussion conditions among Iranian EFL learners. Accordingly, the study addresses the following research questions:

Research question 1: Do EFL learners participating in prompt-supported classroom discussions demonstrate significantly higher posttest engagement than learners participating in conventional discussions?

Research question 2: Do EFL learners participating in prompt-supported classroom discussions demonstrate significantly higher posttest CTD than learners participating in conventional discussions?

3. METHOD

Participants

This study involved 50 female EFL learners recruited through convenience sampling from an English language institute in Ahvaz, Iran. Initially, 70 learners from the accessible population at the institute were screened for eligibility. The Oxford Quick Placement Test (OQPT) was administered to establish a relatively homogeneous proficiency group, and 50 learners whose scores fell within one standard deviation above or below the screening-sample mean were retained. This criterion was used for sample homogeneity rather than as a stand-alone external definition of intermediate proficiency; the participants were classified as intermediate on the basis of their OQPT performance and the placement context of institute. The participants ranged in age from 16 to 22 years. Following screening, the 50 eligible learners were randomly assigned in equal numbers to an experimental group ($n = 25$) and a control group ($n = 25$). Random assignment was intended to reduce systematic pre-intervention differences between the conditions while maintaining broadly comparable English proficiency.

A sensitivity power analysis using G*Power (Faul et al., 2007) was conducted to determine the effect sizes detectable with the final sample. With 25 participants in each group, a two-tailed independent-samples t test at $\alpha = .05$ provided 80% power to detect a between-group standardized effect of approximately $d = 0.81$. For within-group pretest-posttest comparisons, a sample of 25 participants provided 80% power to detect an effect of approximately $d_z = 0.58$. The design had 80% power to detect moderate-to-large within-group effects and large between-group effects; smaller effects may have gone undetected. Participation was voluntary, and informed consent was obtained before data collection. Confidentiality was maintained by assigning numerical codes in place of personal identifiers. Participants were also informed that they could withdraw from the study at any stage without penalty or adverse consequences.

Instruments

Oxford Quick Placement Test

The OQPT, developed by Oxford University Press and the University of Cambridge Local Examinations Syndicate (2001), was used to determine the participants' general English proficiency and identify a relatively homogeneous group of intermediate-level learners. The 60-item test consists of two parts, with 40 items in Part 1 and 20 more demanding items in Part 2, and assesses contextualized knowledge of grammar, vocabulary, and general language use rather than separate subscales. A typical item requires learners to select the most appropriate completion of a

sentence or short exchange. Because the OQPT is designed to assess English proficiency, the test items were administered in English and were not translated into Persian; however, procedural instructions were clarified in the participants' first language (L1) when necessary. Each response was scored as correct or incorrect, and learners whose scores fell within one standard deviation above or below the screening-sample mean were retained as intermediate-level participants. The standardized content of the test was preserved without modification. Previous research has reported satisfactory reliability and validity for the OQPT, with reliability estimates approaching .90 (Geranpayeh, 2003). A pilot administration yielded a Kuder–Richardson Formula 20 (KR-20) coefficient of .88, indicating satisfactory internal consistency for proficiency screening.

Engagement Questionnaire

Learner engagement was measured using an adapted version of the Social Media Engagement Scale for Adolescents (SMES-A) developed by Ni et al. (2020). The instrument was selected because its multidimensional structure maps conceptually onto behavioral, cognitive, and affective forms of engagement relevant to classroom discussion. The validated 11-item scale comprises behavioral engagement (4 items), cognitive engagement (3 items), and affective engagement (4 items). For this study, the items were adapted from the social-media context to face-to-face EFL group discussion, and the adapted version is, therefore, treated as a discussion-specific measure rather than as an unchanged application of the original instrument. For example, a behavioral item was reformulated as *"I actively participate in classroom group discussions when I have an opportunity to contribute."* Responses were provided on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating stronger discussion-specific engagement. Before the main study, the questionnaire was translated into Persian through forward translation and back-translation and reviewed by specialists in applied linguistics and educational measurement for clarity, relevance, and conceptual correspondence. It was, then, pilot-tested with learners similar to the target participants. The pilot analysis produced Cronbach's alpha coefficients of .86 for the total scale, .82 for behavioral engagement, .79 for cognitive engagement, and .84 for affective engagement, indicating satisfactory internal consistency. Expert review supported the content relevance and appropriateness of the adapted version, although this local evidence should not be interpreted as automatically transferring all psychometric properties of the original SMES-A to the classroom-adapted form.

Critical Thinking Disposition Scale

The critical-thinking outcome was operationalized as CTD and assessed using the 11-item Critical Thinking Disposition Scale (CTDS) developed by Sosu (2013). The CTDS was selected because it captures the learners' openness to alternative perspectives and their tendency to question evidence and reflect on judgments, dispositional dimensions directly relevant to discussion-based learning. The two-dimensional structure used in the present study corresponds to Sosu's CTDS and includes Critical Openness (7 items) and Reflective Scepticism (4 items). The items were adapted minimally to fit the classroom-discussion context while retaining their original conceptual focus. For instance, an illustrative Critical Openness item was phrased as *"During classroom discussions, I try to consider the broader implications of the ideas being discussed."* The participants responded on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating a stronger CTD. Prior to the main study, the scale was translated into Persian through forward translation and back-translation and evaluated by experts in applied linguistics, educational psychology, and measurement for conceptual equivalence,

clarity, and cultural suitability. A pilot study with 30 learners whose profiles were similar to those of the target participants yielded Cronbach's alpha coefficients of .87 for the total CTDS, .85 for Critical Openness, and .80 for Reflective Scepticism, supporting satisfactory internal consistency. Expert judgments further supported the content validity and contextual appropriateness of the adapted Persian version, which retained the conceptual structure of [Sosu's \(2013\)](#) CTDS.

Data Collection Procedures

Data collection extended over 10 weeks and followed three sequential phases: proficiency screening and pretesting, the 8-week instructional intervention, and posttesting. At the outset, the OQPT was administered to the initial pool of learners to determine their general English proficiency. Learners who met the intermediate-level criterion were retained, after which the 50 eligible participants were randomly assigned to the experimental and control groups ($n = 25$ each). Before the intervention, both groups completed the adapted engagement questionnaire and the CTDS under standardized classroom conditions. These measures established baseline engagement and CTD scores for subsequent within- and between-group comparisons.

During the 8-week intervention, both groups attended regular face-to-face classes and discussed parallel topics; however, the experimental group completed a structured sequence of prompt-supported activities. In the first intervention session, the instructor explained the purpose of structured prompts and modeled how prompts could move discussion beyond brief or descriptive responses. Across subsequent sessions, prompts were sequenced from relatively guided questions toward more independent analytical and evaluative discussion. The instructor-designed prompts functioned as discussion scaffolds rather than answer templates, leaving learners responsible for generating, defending, and revising their own ideas.

Each experimental-group discussion followed a recurring activity cycle. Learners first considered the topic individually and formulated an initial response, then shared their views in small groups and elaborated on the reasoning or evidence supporting them. During subsequent peer interaction, learners compared viewpoints, identified areas of agreement and disagreement, considered alternative interpretations, and collaboratively evaluated possible solutions. Later in the discussion, learners summarized the strongest arguments and reflected on whether and why their initial position had changed. This cycle was intended to sustain participation and create repeated opportunities for analytical and reflective engagement.

Scaffolding was progressively reduced across the intervention. In earlier sessions, the instructor supplied more explicit follow-up questions and modeled how to request clarification or justification. As learners became familiar with the routine, they assumed greater responsibility for questioning peers, extending previous contributions, and challenging unsupported claims respectfully. The instructor monitored interaction, clarified task requirements, and encouraged balanced participation, intervening when responses remained brief or when discussion was dominated by a small number of learners. This role was intended to facilitate interaction without determining learners' conclusions.

The control group participated in similar face-to-face group discussion activities and worked with matched content, but discussions proceeded according to the conventional classroom approach. Learners were presented with the discussion topics and were invited to exchange opinions without the systematic sequence of prompts, follow-up questioning routines, perspective-taking tasks, counterargument activities, or structured reflection used in the experimental condition. Maintaining matched instructional content and discussion format across the two groups helped ensure that structured prompting remained the principal instructional difference between the conditions.

At the conclusion of the 8-week intervention, the engagement questionnaire and CTDS were administered again to both groups as posttests under conditions equivalent to those used at pretesting. Participant codes were used to match pretest and posttest responses while preserving confidentiality.

Data Analysis

The data were analyzed using IBM SPSS Statistics (Version 28). Descriptive statistics, including means and standard deviations, summarized engagement and CTD scores, and inferential statistics addressed the two research questions. Independent-samples *t* tests were used for the primary between-group comparisons because the research questions concerned whether the experimental and control groups differed at posttest; pretest independent-samples *t* tests were also used to examine baseline differences. Paired-samples *t* tests were additionally conducted to describe within-group change over time. Before the between-group analyses, the assumptions for parametric testing were examined. Independence of observations was satisfied by the study design, the Kolmogorov-Smirnov tests were nonsignificant across the measurement conditions (all $p > .05$), and Levene’s tests supported homogeneity of variance for engagement and CTD at pretest ($p = .447$ and $p = .426$, respectively) and posttest ($p = .646$ and $p = .191$, respectively). The groups also showed no significant pretest differences in engagement ($p = .686$) or CTD ($p = .681$), indicating broadly comparable baseline scores. All tests were two-tailed, with statistical significance set at $\alpha = .05$.

4. RESULTS

The results are presented in relation to the two research questions concerning between-group differences in posttest engagement and CTD. Descriptive statistics are reported first, followed by between-group and within-group inferential comparisons.

Table 1: Descriptive Statistics for Pretest Engagement and Critical Thinking Disposition Scores by Group

Measure	Group	<i>n</i>	<i>M</i>	<i>SD</i>
Engagement	Experimental	25	33.00	7.36
	Control	25	33.80	6.53
Critical thinking disposition	Experimental	25	27.00	7.36
	Control	25	26.08	8.34

Note. Higher scores indicate higher levels of the respective construct.

At pretest, the groups displayed similar patterns for both engagement and CTD, providing a descriptive indication of broadly comparable baseline performance.

Independent-samples *t* tests were conducted to determine whether the groups differed significantly at baseline. The baseline tests in Table 2 showed no statistically significant difference for engagement, $t(48) = -0.41$, $p = .686$, Cohen’s $d = -0.11$, or CTD, $t(48) = 0.41$, $p = .681$, Cohen’s $d = 0.12$. The effect sizes were very small, indicating little evidence of meaningful baseline differences between the groups.

Table 2: Independent-Samples t-Tests Comparing Experimental and Control Groups at Pretest

Measure	Mean Difference	<i>t</i>	df	<i>p</i>	Cohen's <i>d</i>
Engagement	-0.80	-0.41	48	.686	-0.11
Critical thinking disposition	0.92	0.41	48	.681	0.12

Note. Mean difference = experimental group minus control group.

Table 3 summarizes the posttest descriptive statistics. The experimental group showed higher mean scores than the control group for both engagement and CTD, a pattern examined inferentially in the subsequent between-group tests.

Table 3: Descriptive Statistics for Posttest Engagement and Critical Thinking Disposition Scores by Group

Measure	Group	<i>n</i>	<i>M</i>	<i>SD</i>
Engagement	Experimental	25	48.20	11.00
	Control	25	41.00	13.29
Critical thinking disposition	Experimental	25	38.60	7.29
	Control	25	30.88	9.65

Note. Higher scores indicate higher levels of engagement and critical thinking disposition.

The posttest comparisons directly addressed the two research questions. Table 4 shows that the experimental group scored significantly higher than the control group on engagement, $t(48) = 2.09$, $p = .042$, Cohen's $d = 0.59$, indicating a moderate between-group effect. The experimental group also scored significantly higher on CTD, $t(48) = 3.19$, $p = .002$, Cohen's $d = 0.90$, representing a large between-group effect.

Table 4: Independent-Samples t-Tests Comparing Experimental and Control Groups at Posttest

Measure	Mean Difference	<i>t</i>	df	<i>p</i>	Cohen's <i>d</i>
Engagement	7.20	2.09	48	.042	0.59
Critical thinking disposition	7.72	3.19	48	.002	0.90

Note. Mean difference = experimental group minus control group. Positive values indicate higher scores for the experimental group.

Paired-samples t tests were also conducted to describe pretest-to-posttest changes within each condition. Engagement and CTD increased significantly in both groups. These within-group tests indicate change over time, but do not establish that the magnitude of change differed significantly between groups; the posttest between-group comparisons in Table 4 provide the primary evidence relevant to the two research questions.

Table 5: Paired-Samples t-Tests for Pretest-to-Posttest Changes Within Groups

Measure	Group	Mean Change	SD of Change	<i>t</i>	df	<i>p</i>	Cohen's <i>d_z</i>
Engagement	Experimental	15.20	11.22	6.77	24	< .001	1.35
Critical thinking disposition	Experimental	11.60	10.68	5.43	24	< .001	1.09
Engagement	Control	7.20	13.70	2.63	24	.015	0.53
Critical thinking disposition	Control	4.80	8.23	2.92	24	.008	0.58

Note. Mean change = posttest minus pretest. Positive values indicate improvement. Cohen's *d_z* represents the standardized mean change for paired observations.

5. DISCUSSION

This study examined whether systematically designed prompts could support Iranian EFL learners’ engagement and CTD during face-to-face group discussions. Learners in the prompt-supported condition demonstrated higher posttest engagement and CTD than those in the conventional discussion condition, a pattern consistent with structured prompts providing clearer interactional direction while creating repeated opportunities for learners to explain, justify, compare, evaluate, and reconsider ideas in English. The advantage of the prompt-supported condition was evident for both outcomes, although the standardized between-group effect was numerically larger for CTD than for engagement. One interpretation is that this difference reflects the close correspondence between the intervention and the reflective and evaluative processes represented by CTD, whereas engagement is likely to be influenced by a broader range of motivational, interpersonal, and contextual factors. Importantly, both groups improved from the pretest to posttest, indicating that the participation in the face-to-face group discussion itself can provide valuable opportunities for interaction and reflection. The higher posttest scores in the prompt-supported condition, therefore, point to an additional benefit associated with systematically organizing discussion rather than implying that conventional group discussion was ineffective.

The engagement finding is consistent with classroom questioning research showing that higher-level questions can deepen participation and critical dialogue in English-language classrooms (DeWaelshche, 2015). The present findings extend this line of research by suggesting that the value of prompting may lie not only in the cognitive level of individual questions, but also in the way prompts are sequenced throughout an interaction. Rather than merely asking learners to state opinions, the intervention repeatedly required them to elaborate, respond to peers, consider alternative viewpoints, justify claims, and revisit previous contributions. One possible mechanism is reduced uncertainty about what constituted an appropriate contribution, with the prompt structure offering learners clearer entry points into the developing discussion. This form of guidance is potentially useful in EFL contexts because learners must simultaneously manage linguistic formulation, comprehension of peers’ contributions, turn-taking, and development of ideas. The observed pattern is consistent with the prompts helping learners devote greater attention to meaningful participation by making the interactional expectations of the task more transparent. Evidence from AI-mediated environments provides an indirect parallel, as Hwang et al. (2025), Deguchi et al. (2026), and Bozorgian and Rahimi (2025) similarly indicate that purposeful interaction design can influence the quality of learner engagement. Although those studies were

conducted in technologically different contexts, the present findings suggest that the broader pedagogical value of carefully structured prompting can also extend to learner-to-learner classroom interaction without requiring AI-mediated responses.

A further explanation for the engagement result concerns the social organization of participation. Conventional discussion tasks can create opportunities for learners to speak without necessarily ensuring that their contributions become meaningfully interconnected. Learners can state parallel opinions while paying limited attention to the reasoning of others. In contrast, the prompt-supported activities incorporated follow-up questions, requests for justification, perspective-taking, counterargument, and responses to peers, potentially encouraging greater reciprocity. Learners were, therefore, expected not only to express their own views, but also to engage with ideas emerging from the group. This continuing requirement to attend to and build on peers' contributions could sustain involvement beyond an initial speaking turn. [Bandura's \(1986\)](#) SCT offers additional support for this interpretation because learners could observe how peers interpreted questions, defended positions, responded to disagreement, and reformulated arguments. This structure could also have made these practices more visible and provided models that learners could subsequently adopt or adapt in their own participation. Thus, the engagement advantage is consistent with a combination of clearer interactional direction, reciprocal participation, and opportunities for peer modeling rather than simply a greater quantity of classroom talk.

The CTD finding is likewise consistent with previous EFL research. [DeWaelche \(2015\)](#) linked higher-order questioning with critical classroom dialogue, while [Bakhtiari Moghadam et al. \(2023\)](#) reported benefits of explicit CT-oriented instruction among Iranian EFL learners. More recent studies also suggest that inquiry and structured questioning can encourage reasoned judgment and deeper engagement with alternative viewpoints ([Nasseh et al., 2026](#); [Shang et al., 2025](#)). A plausible explanation for the comparatively strong CTD result is that the intervention repeatedly required learners to consider alternative perspectives, defend claims, evaluate reasons, respond to disagreement, and reconsider initial positions. These activities closely correspond to the openness and reflective scepticism represented in the CTDS. Repeated engagement in such practices could, therefore, encourage learners to view questioning, reconsideration, and attention to competing perspectives as normal components of classroom participation. The face-to-face nature of the activity could also have contributed to this process because peers introduced interpretations and disagreements that required immediate consideration. In this respect, the present findings differ from [Liang and Wu \(2024\)](#), whose quantitative analysis did not identify a significant improvement in CT scores following a ChatGPT-supported program. Differences in interactional setting, instructional organization, and outcome measurement offer one possible explanation for this contrast. The current study assessed CTD rather than a separate CT ability measure, and its intervention embedded analytical and evaluative prompting repeatedly within collaborative peer discussion. The findings should, therefore, be interpreted as evidence relevant to learners' orientation toward critical engagement rather than as proof of superior development of general critical thinking ability.

The theoretical perspectives underlying the study provide complementary explanations for these patterns. From constructivist and sociocultural perspectives, the pattern is consistent with prompts scaffolding learners' construction and revision of meaning by directing them toward elaboration, comparison, justification, and reconsideration while leaving the substance of their conclusions open ([Piaget, 1973](#); [Vygotsky, 1978](#)). This balance between guidance and intellectual ownership is potentially important because learners received support for how to engage without being told what position to adopt. SDT offers a parallel motivational explanation. Within this account, structured prompts could support competence by clarifying the nature of expected

contributions, autonomy by allowing learners to formulate and defend their own positions, and relatedness by requiring meaningful attention to peers' ideas (Deci & Ryan, 1985). Instructional structure and autonomy are, therefore, not necessarily contradictory; structure can facilitate engagement when it clarifies expectations without controlling learners' conclusions. CLT provides a further explanation. Face-to-face EFL discussion requires learners to understand the topic, formulate language, monitor peers, manage turn-taking, and develop an argument simultaneously. Vague discussion instructions can consume cognitive resources because learners must determine both what to discuss and how to organize a contribution. The more explicit and progressively sequenced prompts could have reduced this avoidable ambiguity and directed attention toward task-relevant reasoning processes (Sweller, 1988). The prompts did not necessarily make the intellectual task easier; rather, their structure could have organized its demands more effectively.

Taken together, the results suggest that the effects of prompt-supported discussion are unlikely to arise from the mere presence of additional questions. The intervention combined cognitive direction, reciprocal interaction, progressive scaffolding, peer modeling, and opportunities for learners to exercise choice over the content of their contributions. Together, these elements could have created conditions in which engagement and critical reflection reinforced one another: sustained participation increased exposure to alternative ideas, while analytical requirements gave that participation greater cognitive substance. This reciprocal process provides a plausible explanation for why both engagement and CTD were higher in the prompt-supported condition.

6. CONCLUSIONS

This study found that the learners participating in prompt-supported face-to-face discussions achieved higher post-intervention engagement and CTD than those participating in conventional discussions. The higher post-intervention scores are consistent with the interpretation that structured prompting can organize classroom interaction in ways that sustain participation and encourage reflective and evaluative engagement with ideas. The study contributes to the literature by extending deliberate prompt design beyond predominantly human-AI interaction and by offering an integrative theoretical interpretation of how structured prompting may coordinate motivational, social, and cognitive processes during EFL discussion. These mechanisms were not measured directly, however, and should be interpreted as theoretically grounded explanations rather than demonstrated causal pathways.

Pedagogical Implications

The findings have several pedagogical implications. EFL teachers can use graduated prompt sequences and targeted follow-up questions to extend brief responses, promote more balanced participation, and elicit justification or consideration of alternatives. At the teacher-education level, prompt-design practice can be incorporated into professional-development programs. For curriculum and materials development, structured questioning routines and brief reflection stages can be embedded within speaking units. At the program level, implementation could be supported through shared prompt repositories and peer observation, while discussion rubrics could be aligned with behaviors such as elaboration, responsiveness to alternative perspectives, and sustained contribution. Over time, learners can gradually assume greater responsibility by generating follow-up questions and requesting clarification or evidence from peers, reducing dependence on teacher-generated scaffolding.

Limitations and Directions for Further Research

The findings should be interpreted in light of several limitations of the present design. All participants were female intermediate-level learners from a single language institute, so the observed patterns cannot be assumed to characterize male learners, mixed-gender groups, other proficiency levels, or other institutional EFL contexts. Future studies should deliberately vary gender composition, proficiency, and instructional setting. The 8-week intervention was limited to relatively short-term instructional change and does not establish whether engagement and CTD gains persist once structured prompting is reduced or removed; delayed posttests and longer interventions would provide a stronger test of persistence and generalization beyond the intervention period. Another limitation concerns the analytical design. Although both groups improved from the pretest to the posttest and the posttest scores were higher for the experimental group, the reported analyses did not directly test a group-by-time interaction. Subsequent research could use mixed-design ANOVA or linear mixed-effects modeling to test differential change directly. ANCOVA could be considered separately for baseline-adjusted posttest comparisons.

The study also relied on questionnaire-based outcome measures and did not include a qualitative component, limiting insight into how learners experienced the intervention or how prompt-supported interaction unfolded in the classroom. A useful extension would combine outcome measures with recorded classroom interaction, discourse analysis, stimulated recall, interviews, or learner reflections. Experimentally comparing different prompt architectures, such as justification, counterargument, evidence evaluation, and perspective-taking prompts, would help identify which components contribute most strongly to engagement and CTD rather than treating structured prompting as a single undifferentiated intervention.

References

- Aizikovitsh-Udi, E., & Cheng, D. (2015). Developing critical thinking skills from dispositions to abilities: Mathematics education from early childhood to high school. *Creative Education*, 6(4), 455–462. <https://doi.org/10.4236/ce.2015.64045>
- Aydın Yıldız, T. (2023). The impact of ChatGPT on language learners' motivation. *Journal of Teacher Education and Lifelong Learning*, 5(2), 582–597. <https://doi.org/10.51535/tell.1314355>
- Bakhtiari Moghadam, Z., Narafshan, M. H., & Tajadini, M. (2023). The effect of implementing a critical thinking intervention program on English language learners' critical thinking, reading comprehension, and classroom climate. *Asian-Pacific Journal of Second and Foreign Language Education*, 8, Article 15. <https://doi.org/10.1186/s40862-023-00188-3>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bozorgian, H., & Rahimi, H. (2025). Peer e-feedback and ChatGPT-4o in EFL writing: A cognitive-interpersonal comparison based on EFL students. *Journal of Language and Education*, 11(4), 54–65. <https://doi.org/10.17323/jle.2025.27195>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- Deguchi, A., Tsuji, H., & Kitazawa, T. (2026). Designing epistemic partnerships with generative AI: A four-stage model of student engagement in scientific argumentation. *Research in Science Education*. Advance online publication. <https://doi.org/10.1007/s11165-026-10370-5>

- DeWaelsche, S. A. (2015). Critical thinking, questioning and student engagement in Korean university English courses. *Linguistics and Education*, 32, 131–147.
<https://doi.org/10.1016/j.linged.2015.10.003>
- Ding, N., Hu, S., Zhao, W., Chen, Y., Liu, Z., Zheng, H.-T., & Sun, M. (2022). OpenPrompt: An open-source framework for prompt-learning. In V. Basile, Z. Kozareva, & S. Stajner (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 105–113). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-demo.10>
- Ebadi, S., Zandi, M., & Fakher Ajabshir, Z. (2024). The effect of language learning engagement on L2 motivation over Instagram. *Acta Psychologica*, 250, Article 104512.
<https://doi.org/10.1016/j.actpsy.2024.104512>
- Eerdemutu, L., Dewaele, J.-M., & Wang, J. (2024). Developing a short language classroom engagement scale (LCES) and linking it with needs satisfaction and achievement. *System*, 120, Article 103189. <https://doi.org/10.1016/j.system.2023.103189>
- Erawan, P. (2010). Developing life skills scale for high school students through mixed methods research. *European Journal of Scientific Research*, 47(2), 169–186.
- Farsinejad, M., & Haddad Narafshan, M. (2024). The effect of watching sustainability-focused animations on English language learners' critical thinking and reading comprehension. *Interdisciplinary Studies in English Language Teaching*, 2(2), 16–30.
<https://doi.org/10.22080/iselt.2024.5137>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Florea, N. M., & Hurjui, E. (2015). Critical thinking in elementary school children. *Procedia - Social and Behavioral Sciences*, 180, 565–572.
<https://doi.org/10.1016/j.sbspro.2015.02.161>
- 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>
- Gastineau, J. (2024). *Exploring the impact of generative AI prompt engineering in higher education: A study in undergraduate and graduate business analytics courses* [Doctoral dissertation, University of Arkansas]. ScholarWorks@UARK.
<https://scholarworks.uark.edu/etd/5551>
- Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *Research Notes*, 12, 8–10.
- Gever, V. C., Tunca, E. A., Boluwatife, A. A., Nwogbo, V. C., Chinweobo-Onuoha, B. N., Ugwuoke, J. C., & Talabi, F. O. (2021). Visual media and learning: Effect of interactive television instruction as an intervention strategy for improving the critical thinking skills and disposition of out-of-school nomadic children in Nigeria. *Learning and Motivation*, 76, Article 101767. <https://doi.org/10.1016/j.lmot.2021.101767>
- Grabb, D. (2023). The impact of prompt engineering in large language model performance: A psychiatric example. *Journal of Medical Artificial Intelligence*, 6, Article 20.
<https://doi.org/10.21037/jmai-23-71>
- Hayes, K. D., & Devitt, A. A. (2008). Classroom discussions with student-led feedback: A useful activity to enhance development of critical thinking skills. *Journal of Food Science Education*, 7(4), 65–68. <https://doi.org/10.1111/j.1541-4329.2008.00054.x>

- Hoorijani, S., & Heidari Tabrizi, H. (2024). Improving EFL learners' argumentative writing through critical thinking disposition: Focus on gender-related differences. *Interdisciplinary Studies in English Language Teaching*, 2(1), 157–174. <https://doi.org/10.22080/iselt.2023.25923.1056>
- Hwang, M., Lee, K.-H., & Lee, H.-K. (2025). A word to the wise: Crafting impactful prompts for ChatGPT. *System*, 133, Article 103756. <https://doi.org/10.1016/j.system.2025.103756>
- Jacobsen, L. J., & Weber, K. E. (2025). The promises and pitfalls of large language models as feedback providers: A study of prompt engineering and the quality of AI-driven feedback. *AI*, 6(2), Article 35. <https://doi.org/10.3390/ai6020035>
- Jones, A., Hook, M., Podduturi, P., McKeen, H., Beitzell, E., & Liss, M. (2022). Mindfulness as a mediator in the relationship between social media engagement and depression in young adults. *Personality and Individual Differences*, 185, Article 111284. <https://doi.org/10.1016/j.paid.2021.111284>
- Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, 38(5), 758–773. <https://doi.org/10.1080/03075079.2011.598505>
- Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development*, 37(1), 58–71. <https://doi.org/10.1080/07294360.2017.1344197>
- Khalili, T., Ketabi, S., Pnevmatikos, D., & Mahabadi, D. N. A. (2024). Critical thinking and L2 learning in the Persian context: A meta-analysis. *Thinking Skills and Creativity*, 54, Article 101640. <https://doi.org/10.1016/j.tsc.2024.101640>
- Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, 6, Article 100225. <https://doi.org/10.1016/j.caeai.2024.100225>
- Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, 22, Article 7. <https://doi.org/10.1186/s41239-025-00503-7>
- Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, 54, Article 101645. <https://doi.org/10.1016/j.tsc.2024.101645>
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), Article 195, 1–35. <https://doi.org/10.1145/3560815>
- Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (Article 384, pp. 1–23). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3501825>
- Luo, N., Wang, Y., Zhang, Z. (V.), Zhou, Y., & Zhao, R. (2026). Impact of prompt sophistication on ChatGPT's output for automated written corrective feedback. *ReCALL*, 38(2), 240–254. <https://doi.org/10.1017/S095834402510044X>
- Mzwri, K., & Turcsányi-Szabó, M. (2025). The impact of prompt engineering and a generative AI-driven tool on autonomous learning: A case study. *Education Sciences*, 15(2), Article 199. <https://doi.org/10.3390/educsci15020199>
- Nasseh, A., Zareian, G., Ghaniabadi, S., & Amirian, S. M. R. (2026). An inquiry method for critical thinking instruction in the Iranian EFL context. *Thinking Skills and Creativity*, 61, Article 102160. <https://doi.org/10.1016/j.tsc.2026.102160>
- Ni, X., Shao, X., Geng, Y., Qu, R., Niu, G., & Wang, Y. (2020). Development of the social media engagement scale for adolescents. *Frontiers in Psychology*, 11, Article 701.

- <https://doi.org/10.3389/fpsyg.2020.00701>
- Oxford University Press, & University of Cambridge Local Examinations Syndicate. (2001). *Quick Placement Test*. Oxford University Press.
- Pescatore, C. (2007). Current events as empowering literacy: For English and social studies teachers. *Journal of Adolescent & Adult Literacy*, 51(4), 326–339.
<https://doi.org/10.1598/JAAL.51.4.4>
- Piaget, J. (1973). *To understand is to invent: The future of education*. Viking Press.
- Polat, O., & Aydin, E. (2020). The effect of mind mapping on young children's critical thinking skills. *Thinking Skills and Creativity*, 38, Article 100743.
<https://doi.org/10.1016/j.tsc.2020.100743>
- Rahimi, H., Pishghadam, R., Shayesteh, S., & Al Abdwani, T. (2026). The interplay of teacher energy, students' active and passive motivation, and their willingness to attend classes. *Language Teaching Research Quarterly*, 54, 212–239.
<https://doi.org/10.32038/ltrq.2026.54.09>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, Article 101860.
<https://doi.org/10.1016/j.cedpsych.2020.101860>
- Sang, Y., & Hiver, P. (2021). Engagement and companion constructs in language learning: Conceptualizing learners' involvement in the L2 classroom. In P. Hiver, A. H. Al-Hoorie, & S. Mercer (Eds.), *Engagement in the language classroom* (pp. 17–38). Multilingual Matters.
- Schmidt-Fajlik, R. (2023). ChatGPT as a grammar checker for Japanese English language learners: A comparison with Grammarly and ProWritingAid. *AsiaCALL Online Journal*, 14(1), 105–119. <https://doi.org/10.54855/acoj.231417>
- Shang, J., Huang, Y., Xu, M., Huang, Y., Shen, X., Wang, G., Wang, Y., & Zhang, L. (2025). Competition between human learners and ChatGPT: Enhancing university EFL students' reading comprehension and critical thinking through competitive questioning. *Computer Assisted Language Learning*. Advance online publication.
<https://doi.org/10.1080/09588221.2025.2577356>
- Sosu, E. M. (2013). The development and psychometric validation of a Critical Thinking Disposition Scale. *Thinking Skills and Creativity*, 9, 107–119.
<https://doi.org/10.1016/j.tsc.2012.09.002>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292.
<https://doi.org/10.1007/s10648-019-09465-5>
- Tavassoli, K., Bahrami, M., & Ghamoushi, M. (2024). The effect of the six thinking hats technique on Iranian EFL learners' writing skill and their attitudes toward it. *Interdisciplinary Studies in English Language Teaching*, 2(2), 67–88.
<https://doi.org/10.22080/iselt.2025.27524.1073>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Woo, D. J., Wang, D., Yung, T., & Guo, K. (2026). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge and prompt engineering ability:

- A mixed methods study. *British Educational Research Journal*, 52(2), 1442–1469. <https://doi.org/10.1002/berj.70087>
- Wu, T.-T., Hapsari, I. P., & Huang, Y.-M. (2026). Intelligent chatbot-supported collaborative learning: Impact on student engagement and English speaking skills. *ReCALL*, 38(2), 255–273. <https://doi.org/10.1017/S0958344025100451>
- Zou’bi, R. A. (2021). The impact of media and information literacy on acquiring the critical thinking skill by the educational faculty’s students. *Thinking Skills and Creativity*, 39, Article 100782. <https://doi.org/10.1016/j.tsc.2020.100782>